Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
Department of Psychology, University of Michigan
Ann Arbor, Michigan

DAVID A. GRANT, Associate Editor
University of Wisconsin

WILLIAM K. ESTES, Associate Editor
Indiana University

Consulting Editors

NORMAN H. ANDERSON, *University of California, Los Angeles*
E. JAMES ARCHER, *University of Wisconsin*
FRED ATTNEAVE, III, *University of Oregon*
CLETUS J. BURKE, *Indiana University*
JAMES DEESE, *Johns Hopkins University*
PAUL M. FITTS, *University of Michigan*
FRANK A. GELDARD, *University of Virginia*
JAMES J. GIBSON, *Cornell University*
HAROLD W. HAKE, *University of Illinois*
ARTHUR L. IRION, *Tulane University*
HOWARD H. KENDLER, *New York University*
HERSCHEL W. LEIBOWITZ, *International Business Machines Corporation, Bethesda, Maryland*
DONALD B. LINDSLEY, *University of California, Los Angeles*
KENNETH MACCORQUODALE, *University of Minnesota*
QUINN MCNEMAR, *Stanford University*
LEO POSTMAN, *University of California, Berkeley*
L. STARLING REID, *University of Virginia*
KENNETH W. SPENCE, *State University of Iowa*
DELOS D. WICKENS, *Ohio State University*

CONTENTS

Some Variables Influencing the Rate of Gain of Information: R. W. BRAINARD, T. S. IRBY, P. M. FITTS, AND E. A. ALLUISI 105
The Ability to Filter Noise from a Visual Task when the Noise and Signal are Presented Sequentially: D. E. ERLICK 111
Perception of the Most Frequent Category of a Random Series as a Function of the Number of Categories: D. E. ERLICK 115
Manifest Anxiety Scale Score and the Ready Signal in Classical Conditioning: W. F. PROKASY, JR. AND F. L. WHALEY 119
A Test of Two Interpretations of the Apparent Size Effects in a Distorted Room: W. EPSTEIN 124
Confusion Matrix Analysis of Single and Multidimensional Auditory Displays: M. H. HODGE AND I. POLLACK 129
Reaction Time to Stimuli Masked by Metacontrast: E. FEHRER AND D. RAAB 143
The Effect of Postural Factors on the Distribution of Tactual Sensitivity and the Organization of Tactual-Kinaesthetic Space: J. H. MCFARLAND, H. WERNER, AND S. WAPNER 148
Acquisition and Transfer in Partial Reinforcement: E. J. CAPALDI AND M. G. SENKO 155
Amount, Delay, and Position of Delay of Reinforcement as Parameters of Runway Performance: E. R. WIST 160
The Influence of Early Experience on the Frustration Effect: A. AMSEL AND E. C. PENICK ... 167
The Progression-Regression Hypotheses in Perceptual-Motor Skill Learning: A. H. FUCHS ... 177
The Competence of the Model and the Learning of Imitation and Nonimitation: M. E. ROSENBAUM AND I. F. TUCKER 183
Effects of Mental Set upon Vestibular Nystagmus: W. E. COLLINS 191
Response to Simultaneous Stimulation of Two Sense Modalities: J. A. ADAMS AND R. W. CHAMBERS 198
Nonreinforcement and Neutralization of Stimuli: D. LABERGE, J. G. GREENO, AND O. F. PETERSON 207
Supplementary Report: Partial Reinforcement and Amount of Reinforcement as Determinants of Instrumental Licking Rates: S. H. HULSE AND W. E. BACON 214
Supplementary Report: Shock Intensity and Unconditioned Responding in a Shuttle Box: T. R. TRABASSO AND R. W. THOMPSON 215

American Psychological Association

Vol. 63, No. 2 February 1962

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc., at Prince and Lemon Sts., Lancaster, Pa. The subscription rate per volume is $10.00, or $20.00 annually. Single copies are $2.00. Subscriptions, orders, address changes, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This Journal publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects, except when the latter are specifically oriented toward the extension of behavior theory. Experimental psychometric studies and studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology.

Normally, articles exceeding 20 printed pages cannot be accepted. Within this limit, the piecemeal experiment-by-experiment reporting of psychological research is discouraged, and the reporting of the data and conclusions of substantial segments of research efforts is encouraged. Specifically, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) *must* be presented in a single article. Special provision is made for the publication of Supplementary Reports and Replication Reports in articles of not more than 1.5 printed pages (see Editor's Note, this Journal, 1957, **53**, 1–2).

Manuscripts *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc., as described in the *Publication Manual of the American Psychological Association.* When in doubt about practices of the Journal, authors should examine a recent issue.

Articles are published in the order of their receipt, except in rare circumstances. Authors are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author. Priority in publication is given to articles whose authors assume the full cost of publication. In 1961 the publication cost is $30 per printed page.

All **regular** articles, **supplementary reports,** and **replication reports** should be submitted to the Editor-elect: David A. Grant: Department of Psychology, University of Wisconsin; Madison 6, Wisconsin. **Early publication** articles should be sent to the Editor: Arthur W. Melton; Department of Psychology, University of Michigan; Ann Arbor, Michigan until *July 1, 1962.* Thereafter they should be sent to the Editor-elect. Manuscripts must be accompanied by an abstract of 100 to 120 words typed on a separate sheet of paper. The abstract should conform to the style of *Psychological Abstracts.* Detailed instructions for preparation of the abstracts may be obtained from the Editor or from the APA Central Office.

HELEN ORR
Acting Production Manager

Second-class postage paid at Lancaster, Pa.

Send address changes to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

© 1962 by the American Psychological Association, Inc.

Journal of
Experimental Psychology

VOL. 63, NO. 2 FEBRUARY 1962

SOME VARIABLES INFLUENCING THE RATE OF GAIN OF INFORMATION [1]

ROBERT W. BRAINARD, THOMAS S. IRBY,[2] PAUL M. FITTS,[3] AND EARL A. ALLUISI [4]

Ohio State University

Choice time, or disjunctive reaction time (RT), has been found to be a linear increasing function of the average amount of information transmitted (H_t) per stimulus-response event in a variety of tasks (e.g., Gregg, 1954; Hick, 1952; Hyman, 1953). Hick found that the same function fitted the data when errorless performance was required and also when *S* speeded up his responses to the point where a substantial number of errors occurred. Hyman, demanding errorless performance, found that RT was a linear function of stimulus information, H_s, (and, therefore, of H_t) when three different procedures for varying average H_s were used: (*a*) the number of equally probable alternative stimuli was varied, (*b*) the relative frequency of occurrence of different stimuli was varied, and (*c*) the degree of sequential dependence of stimuli was varied. The mathematical equivalence of these three methods of varying stimulus uncertainty has been shown by Shannon and Weaver (1949); their psychological equivalence can be inferred from the results of Hyman's study.

Similar results have been obtained with the use of self-paced serial-reaction tasks. For example, in a serial card-sorting task Crossman (1953) found choice time to be linearly related to the average information processed per response, and in a serial switch-moving task Archer (1954) found that the average time per response increased as a linear function of the relevant information presented in the stimulus (over the range from 1 to 4 bits/stimulus), but was independent of the irrelevant information

[1] This report is a consolidation of data from two master's theses by Brainard and Irby; these were completed in 1956 and 1957 as part of a general research program for which P. M. Fitts and E. A. Alluisi were responsible, and which was supported in part by the United States Air Force under Contract No. AF 33(616)-43 and -3612, with the Ohio State University Research Foundation, monitored by the Aerospace Medical Laboratory, Wright Air Development Command. Permission is granted for reproduction, translation, publication, use, and disposal in whole or in part by or for the United States Government.

[2] Now at Dunlap & Associates, Incorporated, Stamford, Connecticut.

[3] Now at the University of Michigan.

[4] Now at Emory University.

presented (over a range from 0 to 2 bits/stimulus).

Although a linear relation between RT and H_t of the form $RT = a + bH_t$ has been consistently obtained, the slopes and intercepts of the functions, i.e., the numerical values of a and b, have varied widely from one task to another. Thus it is clear that parameters other than the informational characteristics of the task are significant in determining RT, specifically in modifying the slope and intercept of the function.

The present study was planned to investigate some of the parameters that may be responsible for the differences in slope and intercept obtained in earlier studies. Specifically, the objective was to determine the effects of (*a*) two types of stimulus codes (numerals and light patterns), (*b*) two types of response codes (vocal and finger), and (*c*) self-paced serial vs. discrete-reaction tasks. The four combinations of stimulus and response codes were chosen to provide high and low levels of S-R compatibility. Three levels of stimulus uncertainty (1, 2, and 3 bits/stimulus) were obtained by varying the number of equally probable alternative stimuli. Only naive *S*s were used, each *S* serving under only one S-R coding condition and one level of stimulus uncertainty. Thus the experiment also provides further tests of the generality of the linear relation of RT to H_t.

Method

The study was conducted in two parts, which were identical except for the specific S-R pairings used. Pairings of (*a*) key pressing responses to spatially arranged lights and (*b*) vocal responses to numerals were used in the first part. It was predicted that these S-R ensembles should show high compatibility i.e., give relatively high values of H_t per second. Pairings of (*c*) key pressing responses to numerals and (*d*) vocal responses to spatially arranged lights were employed in the second part. It was predicted that the latter S-R ensembles would be relatively less compatible since neither physical correspondence nor overlearning operated in its favor. The data of the two parts of the experiment were collected at different times, but the same apparatus and procedures were used throughout and *S*s were drawn from the same population of university students.

Stimuli.—Two sets of stimuli were used: 0.25-in. high Arabic numerals and 0.5-in. diameter ruby lights. The numerals appeared as projections on the center of a 10-in. diameter circular display of opal glass (brightness of 2.5 ft-L). The lights (440 ft-L) were mounted on a vertically placed panel; they were arranged spatially in two semicircles to correspond with the placement of the finger tips. Each of the stimulus surfaces was located approximately 28 in. directly in front of the seated *S*.

The amount of information characterizing a stimulus event was varied by random selection of stimuli from among two, four, or eight equally likely alternatives (1, 2, or 3 bits/stimulus, respectively). One set of conditions consisted of either two, four, or eight Arabic numerals. At the lowest level of stimulus uncertainty the numerals used were 4 and 7; the numerals 3, 4, 7, and 8 were used at the intermediate level, and all the numerals from 2 through 9 were used at the highest level of uncertainty. The other set of conditions consisted of either two, four, or eight ruby-colored lights. For reference purposes the lights, from left to right, were assigned the numbers 2 through 9; the specific lights used in making up the three levels of stimulus uncertainty were those corresponding to the three sets of numeral stimuli. Any light not being used in a specific condition was replaced with a 0.5-in. diameter metal cap that was set flush with the plane of the panel.

Mode of information presentation.—Stimuli were presented both in a self-paced serial-reaction task and in a discrete-reaction task. In the self-paced task, the first stimulus appeared approximately 3 sec. after a "get ready" burst of broad-band noise was presented through the earphones worn by *S*. Each succeeding stimulus appeared 0.15 sec. after *S* had responded to the previous stimulus. The interval of 0.15 sec. represents the electromechanical switching time of the experimental apparatus; it was subtracted from the self-paced response times to yield the RTs. (No corrections were made for the rise time of the incandescent lights, lags of the voice key or time clocks, or differential brightness of the ruby lights and projected numerals.)

Discrete stimuli were presented at the rate of one stimulus every 10 sec. A red circle in the center of the display screen (or a green light in the center of the array of ruby lights) appeared together with a broad-band noise signal to provide a "get ready" indication during the 2 sec. immediately preceding each stimulus presentation. The stimulus remained on for 8 sec.; it was followed by the 2-sec. warning for the next stimulus, etc.

Responses.—For the key pressing task, a bank of 10 keys was arranged horizontally in two semicircles to correspond with a natural placement of the finger tips. These keys were numbered from left to right, 1, 2, 3, . . . 10. Wooden blocks were inserted under those keys for which no corresponding stimulus would appear. The fingers used in the key pressing responses for the three levels of stimulus complexity were the two index fingers (4 and 7) for the 1-bit/stimulus level, the middle and index fingers of both hands (3, 4, 7, and 8) for the 2-bit/stimulus level, and all but the little fingers for the 3-bits/stimulus level. When making these motor responses, *S*'s task was to depress the key that corresponded to the stimulus presented. Specific errors were recorded on an 8×8 electromechanical counter matrix.

In making vocal responses *S*'s task was to identify the stimulus by responding with its numerical label. The responses were monitored and errors recorded by *E*, who had in front of him a list of the series of stimuli being presented to *S*. The *S*'s responses were also tape recorded so that ambiguities in scoring could be resolved by replaying the record.

Procedure.—The study employed a $4 \times 3 \times 2$ factorial design, consisting of the four S-R codes (light and numeral stimuli paired with vocal and key pressing responses), the three levels of stimulus uncertainty (1, 2, and 3 bits/stimulus), and the two task conditions (self-paced and discrete).

Different groups of 20 *S*s were assigned to each of the combinations of the four S-R pairings with the three levels of stimulus uncertainty (12 groups); however each *S* served under both self-paced and discrete-reaction conditions. Each *S* served for five trials lasting somewhat less than an hour in all. Trials 1, 3, and 5 (100 responses per trial) required self-paced serial reactions, whereas Trials 2 and 4 (50 responses per trial) required discrete reactions. The first self-paced trial was given for purposes of task familiarization and data obtained from it were omitted from analysis.

Each *S* was seated in a small experimental booth. Ambient lighting was furnished by a single miniature fluorescent light; brightness contrast was high and all stimuli were quite legible. Broad-band noise (approximately 60 db.) was presented continuously to *S* through their earphones during all self-paced trials in order to mask extraneous sounds. Earphones were also worn during the discrete-reaction trials, but masking noise was presented only during the 2-sec. interval just prior to the presentation of each stimulus.

Subjects.—The *S*s were 240 undergraduate students at the Ohio State University, 156 men and 84 women. In each part of the study *S*s were assigned to one of the six groups used in that part on the basis of order in reporting to *E*, with the restriction that each group contain 13 men and 7 women.

Results

Three measures of performance were computed: (*a*) average reaction time (RT, in sec/response), (*b*) average information transmitted per response (H_t, in bits/response), and (*c*) average rate of information transmission per response (H_t/RT, in bits/sec).

Relation between RT and H_t.—Average RT and average H_t were computed from the 100 discrete reactions of each *S*; similar estimates were also computed from the last 200 self-paced serial reactions of each *S*. Scores for *S*s were then averaged to obtain group estimates. The relation between these average RTs and H_t's were examined for each of the eight functions obtained for the different combinations of the two modes of information presentation (discrete and self-paced reactions) and the four S-R pairings. The least-squares linear equations of best fit for the eight conditions are given in Table 1. For none of the conditions was the deviation from linearity statistically significant ($P > .05$ in each case).

An average Pearson r between RT and H_t was obtained for all *S*s who used a given S-R code under similar (serial or discrete) conditions. For one of the more compatible conditions

(key pressing responses to spatially arranged lights) and for both of the less compatible conditions the correlations were high (the average correlation was +.84). Thus, for these conditions, approximately 70% of the variation in average RT is accounted for by the variations in H_t. However, the average correlation between RT and H_t was only +.34 for the group tested under conditions in which vocal responses were made to numerical stimuli; here the variations in H_t account for only about 10% of the variations in RT.

Effect of mode of information presentation.—The two modes of information presentation (discrete vs. self-paced serial reactions) produced relatively small differences in performance. The linear relations between RT and H_t were nearly identical within each of the four S-R pairings and the differences between the four pairs of discrete and self-paced conditions in terms of the correlations of RT and H_t were in no case statistically significant ($P > .05$ in each case). There were small differences between the discrete and self-paced serial reactions in both RT and H_t, however, as shown in Fig. 1.

TABLE 1

EMPIRICAL LINEAR EQUATIONS OF BEST FIT FOR AVERAGE RT, IN SEC/RESPONSE, AS A FUNCTION OF AVERAGE H_t, IN BITS/RESPONSE

S-R Pairing	Task	Equation
N-M	SP	$RT = .281 + .183(H_t)$
N-M	D	$RT = .317 + .180(H_t)$
N-V	SP	$RT = .485 + .011(H_t)$
N-V	D	$RT = .360 + .061(H_t)$
L-M	SP	$RT = .229 + .111(H_t)$
L-M	D	$RT = .182 + .136(H_t)$
L-V	SP	$RT = .267 + .205(H_t)$
L-V	D	$RT = .335 + .178(H_t)$

Note.—N = Numerals, L = Lights, M = Motor (key pressing), V = Vocal, SP = Self-paced, D = Discrete.

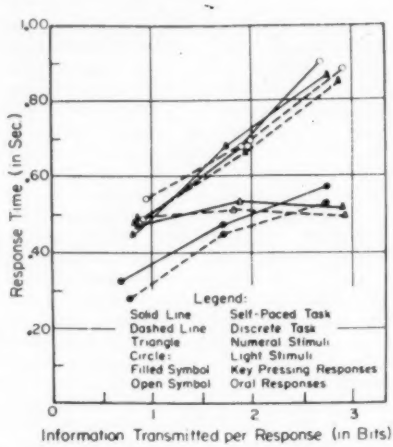

FIG. 1. Reaction time as a function of information transmitted for four S-R codes, and for self-paced vs. discrete tasks.

In the majority of cases the discrete reactions were made slightly slower and more accurately than the self-paced serial reactions; these changes tended to counteract each other, however, since average H_t/sec was almost identical for the two conditions.

S-R compatibility effects.—The most striking S-R compatibility effect is apparent from an examination of the slope constants for the linear equations of Table 1. In all cases the slopes of the functions relating RT and H_t for the less compatible tasks are greater than are those for the most compatible tasks. Thus, differences in RT attributable to the use of different S-R codes are present when only two alternatives are used, but increase markedly as number of alternatives (and H_t per response) increases.

DISCUSSION

The present results agree with those of previous investigations showing that RT is a linear function of H_t. They also agree with previous studies (cf. Alluisi, Muller, & Fitts, 1957) in indicating that average information transmission rates

increase, within limits, as the size of the coding alphabet is increased. These results, it should be noted, are for relatively naive *S*s whereas most previous studies have employed practiced *S*s.

Various forms of linear equations were fitted to the data collected here, including the form proposed by Hick in 1952, $RT = b \log (n + 1)$, where n is the antilogarithm of H_t. The best fit was obtained for the form used in Table 1, $RT = a + b(H_t)$. This mathematical model offers the additional advantage of providing estimates of RT as H_t approaches zero.[5]

Information transmission rate does not appear to be differentially affected by changes from discrete to self-paced serial reactions. There was a tendency for the discrete reactions to be made more accurately but more slowly than the serial; however, these differences in H_t and RT were small and compensating. Thus, the linear equations were essentially identical for each of the four comparisons. It appears, therefore, that the serial task, in which only the .15-sec. machine lag intervened between *S*'s response and the appearance of the next stimulus, permitted as effective information processing as did a rest interval of 8 sec. followed by a 2-sec. warning signal.

Marked differences in the values of the intercept (a) and slope (b) constants were found as a function of the S-R coding combinations employed. In view of the number of previous studies that have obtained similar S-R compatibility effects, including studies comparing vocal and motor responses (Alluisi et al., 1957; Muller, 1955), it is clear that human information processing time is highly dependent on the manner in which stimuli and responses are coded as well as upon the degree of stimulus and response uncertainty. The importance of such compatibility effects is further indicated by results that show these effects to persist even after considerable training (e.g., Fitts & Seeger, 1953) and to become increasingly important as the size of the coding alphabet is increased (as shown here). There is also some evidence that the importance of compatibility effects may increase with age (Suci, Davidoff, & Surwillo, 1960; Welford, 1958), and several writers have proposed that the magnitude of these effects increases as the level of task stress increases.

The results obtained in the vocal-response-to-numerals task are basically different from the other findings. In both the discrete and the self-paced serial task, average RT was only slightly affected by the amount of information transmitted (H_t) per S-R event. As a matter of fact, RTs remained approximately constant for the 2- and the 3-bits/stimulus levels of stimulus uncertainty.

The failure of RT to decrease significantly as the number of alternative stimuli was reduced has been found (a) in studies that required *S* to read words with the number of possible stimulus words varying over a range from 2 to 256 (Pierce & Karlin, 1957; Sumby & Pollack, 1954), (b) in a study wherein *S* was trained for 45,000 trials on a task requiring finger responses to lights (Mowbray & Rhoades, 1959), and (c) in a study in which *S*s made finger responses to vibratory stimuli applied to the finger tips (Leonard, 1959). From a consideration of the tasks used in these studies, the common finding may be given a plausible explanation in terms of the extended learning involved. It seems reasonable to expect that the stronger the habit strength of the S-R combinations required in a specific task, the less steep should be the function relating RT to H_t.

However, in the present study, there

[5] Simple RTs of 5 *S*s (1 male and 4 women) who were members of the laboratory staff were measured with the same apparatus used in this study. Means of 0.316 and 0.440 sec. were obtained for the key pressing and vocal responses, respectively, to the numerals; means of 0.207 and 0.311 sec. were obtained for the same responses to the lights. The σ_m's of these means were 0.008, 0.022, 0.001, and 0.010 sec., respectively. Each mean lies between the discrete and the self-paced serial-reaction intercepts of the appropriate equations of Table 1. The relatively high absolute values for the RTs to numerals may reflect the rise times and lower intensities of the projected numerals; vocal RTs include a slight lag time of the voice key.

is little a priori basis for assuming stronger habit strength for numeral-vocal responses than for light-finger responses, i.e., for predicting less effect of H_t per response for the former type of coding. It appears possible, as an alternative, that the effects of learning may be reflected in the ability to deal effectively with particular sets of stimuli in combination. In the present case the 4–7 and 3–4–7–8 combinations of numerals may have been so incongruent with *S*'s previous experience in dealing with subsets of numerals that he responded as though the entire alphabet of 10 Arabic numerals was being presented. Such a cognitive set presumably would not operate in the case of the lights, since there is no a priori reason to expect a set of any particular size.

Summary

This study investigated the effects on information processing of (*a*) the use of self-paced serial reactions vs. discrete reactions, and (*b*) the use of different stimulus and response codes (both numerals and lights were used as stimuli, and both vocalizations and finger movements were used as responses), and (*c*) the use of three levels of stimulus uncertainty (ranging from 1 to 3 bits/stimulus in unit steps). These several conditions, and the use of naive *S*s, provided further tests of the generality of the function, $RT = a + b(H_t)$.

The results were: (*a*) Reaction time (RT) was an increasing linear function of the average amount of information transmitted (H_t) per stimulus-response event for three of the four S-R pairings employed. The RTs obtained for vocal responses to Arabic numerals, however, were affected only slightly by number of alternatives in the range from 2 to 8. (*b*) The self-paced and discrete tasks gave very similar results. There was a slight tendency for the discrete reactions to be made more accurately but more slowly than the self-paced serial reactions, but the differences both in RT and H_t were small and compensating. The linear functions relating RT with H_t were essentially identical for the two modes of stimulus presentation.

These findings are interpreted as indicating the importance of overlearning in determining S-R compatibility effects. The findings also suggest that *S*'s familiarity in dealing with specific subsets drawn from familiar alphabets may also affect his information-handling rates, when restricted subsets of stimuli and responses are used.

References

Alluisi, E. A., Muller, P. F., Jr., & Fitts, P. M. An information analysis of verbal and motor responses in a forced-paced serial task. *J. exp. Psychol.*, 1957, **53**, 153–158.

Archer, E. J. Identification of visual patterns as a function of information load. *J. exp. Psychol.*, 1954, **48**, 313–317.

Crossman, E. R. F. W. Entropy and choice time: The effect of frequency unbalance on choice-response. *Quart. J. exp. Psychol.*, 1953, **5**, 41–51.

Fitts, P. M., & Seeger, C. M. S-R compatibility: Spatial characteristics of stimulus and response codes. *J. exp. Psychol.*, 1953, **46**, 199–210.

Gregg, L. W. The effect of stimulus complexity on discrimination responses. *J. exp. Psychol.*, 1954, **48**, 289–297.

Hick, W. E. On the rate of gain of information. *Quart. J. exp. Psychol.*, 1952, **4**, 11–26.

Hyman, R. Stimulus information as a determinant of reaction time. *J. exp. Psychol.*, 1953, **45**, 188–196.

Leonard, J. A. Tactual choice reactions: I. *Quart. J. exp. Psychol.*, 1959. **11**, 76–83.

Mowbray, G. H., & Rhoades, M. V. On the reduction of choice reaction times with practice. *Quart. J. exp. Psychol.*, 1959, **11**, 16–23.

Muller, P. F., Jr. Verbalization as a factor in verbal versus motor responses to visual stimuli. Unpublished doctoral dissertation, Ohio State University, 1955.

Pierce, J. R., & Karlin, J. E. Reading rates and the information rate of a human channel. *Bell Sys. tech. J.*, 1957, **36**, 497–516.

Shannon, C., & Weaver, W. *The mathematical theory of communication.* Urbana: Univer. Illinois Press, 1949.

Suci, G. J., Davidoff, M. D., & Surwillo, W. W. Reaction time as a function of stimulus information and age. *J. exp. Psychol.*, 1960, **60**, 242–244.

Sumby, W. H., & Pollack, I. Short-time processing of information. *USAF HFORL tech. Rep.*, 1954, No. 54-6.

Welford, A. T. *Ageing and human skill.* London: Oxford Univer. Press, 1958.

(Received November 17, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 2, 111–114

THE ABILITY TO FILTER NOISE FROM A VISUAL TASK WHEN THE NOISE AND SIGNAL ARE PRESENTED SEQUENTIALLY[1]

DWIGHT E. ERLICK

Aerospace Medical Laboratory, Wright-Patterson Air Force Base

Encoding has proved useful in the processing of excessive information due either to overload conditions or to irrelevance of some stimuli in a situation. In some information processing problems where the information can be categorized, the human processor may handle an overload by recoding the input into another code that contains fewer units or chunks, with more bits per chunk (Miller, 1956). However, when this is not feasible, one may resort to processing only a selective portion of the informational input, i.e., acting as a human information filter. This process of filtering might be viewed as an encoding problem whereby the unused information is encoded into a general category such as noise. Several investigations dealing with irrelevant stimuli might be interpreted in this manner, i.e., the observer acting as a noise filter. The general trend of these studies indicates that simpler tasks tend to be less influenced by noise (Archer, 1954; Green, McGill, & Jenkins, 1953); performance tends to decrease as the quantity of filtering required increases (Archer, Bourne, & Brown, 1955; Gregg, 1954; Hodge, 1959); performance tends to decrease when stimuli to be encoded as noise, i.e., filtered out, have previously been encoded in terms of relevant dimensions of the task (Hodge, 1959); the ability to filter out noise increases as a function of the ease of differentiating or categorizing the relevant from the irrelevant aspects of the task; and practice tends to increase the ability to filter out stimuli encoded as noise (Hodge, 1959; Morin, Forrin, & Archer, 1961). In all the foregoing investigations, the relevant stimuli (signal) and the irrelevant stimuli (noise) were presented in the same modality and occurred in spatial and temporal contiguity.

The purpose of the present experiment was to determine the effect of presenting the signal and noise in the same modality but temporally separated, i.e., sequentially. It might be hypothesized that with a temporal separation of signal and noise, *S* would *have* to attend to each and every unit of noise as it appeared. This procedure might be expected to degrade performance more than when the signal and noise are presented at the same time, a situation where it is more feasible to disregard larger units or chunks of the noise per unit time. It might also be hypothesized that an increase in noise would tend to decrease performance. Based on these hypotheses, a task requiring the sequential analysis of stimuli was selected. The task involved was the judgment of the relative frequency of two random sequential categories. Erlick (1962) has shown that the information transmitted, based on correct identifications of the more frequent category, is linear as a function of the frequency difference between the categories. The task also showed no significant effects due to

[1] The data were collected at the Engineering Psychology Research Project, Antioch College.

practice. The major variable investigated in the following experiment was the effect on this temporal task of varying the amount of noise (irrelevant stimuli) presented in sequential relationship to the task.

METHOD

A punched paper tape reader was used to present a preprogramed random series of different letters which were displayed one at a time in the same location on a 3 × 4 in. screen by means of a rear projection system. The symbols, $3\frac{3}{4}$ in. high, were presented at the rate of 4 per sec. and were on for approximately half of each cycle, producing an equal on-off ratio. This rate and ratio prevented *S* from counting the individual symbols, yet produced a distinct off-time when the same symbol was repeated in sequence. The letters *A* and *B* were the two mutually exclusive categories, i.e., the signal. The 10 letters, *C, D, E, F, G, H, K, L, M, N,* were the irrelevant stimuli, i.e., the noise.

The ratios between the two categories were used: 24:20, and 30:20. These were selected to test the effects of different signal intensities and will be referred to as S_1 and S_2, respectively. On the basis of previous studies (Erlick, 1962) S_1 and S_2 would be predicted to transmit .15 and .49 bits of information, respectively, in terms of identifying the more frequently occurring category.

With S_1 having a total of 44 relevant letters and S_2 having 50 relevant letters, the signal-to-noise ratios as computed in terms of the number of relevant to irrelevant letters were: 1, 2, 4.9, 11, and 22 for S_1; and 1, 2, 5, 10, and 16.7 for S_2. A no-noise condition was also used for both S_1 and S_2.

The interspersing of the noise increases the time period between signal letters. To control for possible interactions between the effects of the noise and the variations in time, the equipment was designed so that the irrelevant letters could be switched off and blank spaces, i.e., empty time intervals, substituted for these letters. In effect, the equipment filtered out the noise, allowing for comparison of performance of the task under conditions with and without noise. Therefore, the interference with the signal was subdivided into the noise conditions (N) and the blank interval conditions (B). A trial required one judgment and consisted of observing a random series of one of the conditions BS_1, NS_1, BS_2, or NS_2 at one of the signal-to-interference ratios or the no-interference condition, which is the same as the no-noise and no-blank interval condition. A session consisted of 16 trials. An experimental period consisted of six sessions each with a different signal-to-interference ratio. The ratios within each of the six sessions were approximately the same, i.e., 1, 2, 4.9; and 5, 10; and 11, 16.7; and 22; and no-interference. All sessions except the no-interference condition were subdivided into a unit of eight B trials and a unit of eight N trials. Within each of these B and N units, S_1 and S_2 were each replicated four times. A separate random selection (with replacement) of the irrelevant letters was made from the 10 irrelevant letters for each of the 4 replicated signal-to-noise ratios within each N unit. Different random sequences of signal and noise letters were also used for each of the four replications. The sequence of S_1 and S_2 within each B and N unit was randomly determined, with a different random sequence being used for each session and each experimental period.

Two practice trials preceded each of the units within each session. Prior to the two practice trials, *E* switched the equipment to either Cond. B or N. The six sessions were presented to six groups of 16 *S*s each, according to a latin square design. The six groups were subdivided so that half of the *S*s received Cond. B first for all sessions within a period and the other half of the *S*s received Cond. N first for all sessions. All 96 *S*s had two experimental periods separated by at least 1 day, with the order of B and N being reversed for the second period. Thus, all *S*s made 8 observations at each of the different experimental conditions for a total of 192 observations. Each experimental period was punched on one continuous tape with an automatic stop programed between trials. The foregoing different randomizations were achieved to some extent by reversing the tapes for half *S*s, thus producing different sequential orders of the *A*s, *B*s, and signal-to-interference ratio. The ratios S_1 and S_2 were counterbalanced so that the letters *A* and *B* each appeared more frequently half the time.

The *S*s were instructed that other letters would appear in some of the series, but their task was to determine only which of the two letters *A* or *B* had appeared more frequently. The same general procedure was used as in previous experiments (Erlick, 1961), including running groups of paid college *S*s and offering a monetary bonus as a function of the number of correct answers.

Results

The following analysis is based on the summated responses of all 96 *S*s. The percentage of correct identifications for each of the two signal intensities (S_1 and S_2) at each of the five signal-to-interference ratios for Cond. B and N and the no-interference condition, were converted into information transmitted. A graphic analysis based on this measure indicated no systematic interactions between preference for the relevant symbols *A* or *B* and any of the other experimental conditions. Also, the amount of information transmitted for the first and second experimental periods was the same, .27 bits, indicating no effects due to practice. Therefore, responses to both sets of symbols and both experimental periods were combined for the following analyses.

The information transmitted at the no-interference condition for signal intensities of S_1 and S_2 was .17 bits and .47 bits, respectively. Comparison of the no-interference condition and the interference conditions averaged over all signal-to-interference ratios for S_1 and S_2 showed the same information transmitted for both conditions at S_2 and a slight decrease (.04 bits) for the interference condition at S_1. The amount of information transmitted averaged over both signal levels and all signal-to-interference ratios was .26 bits for the noise condition and .28 bits for the blank-interval condition. Graphic analysis of the interactions between the noise and blank-interval conditions and signal-to-interference ratios for both signal intensities showed no systematic effects indicating no differences between any of the foregoing conditions (see Fig. 1).

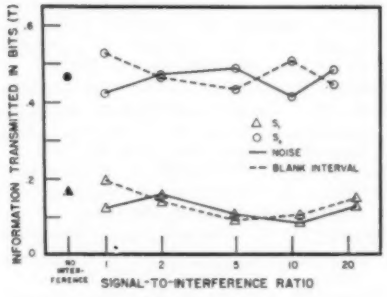

FIG. 1. Information transmitted as a function of the signal-to-interference ratio for noise and blank-interval conditions and for two different signal intensities. (The abscissa is scaled logarithmically. Each point is based on 192 observations except the no-interference condition which is based on 384 observations.)

Discussion

The lack of any systematic bias by the group for the symbols *A* and *B* confirms the results of earlier studies by the author (Erlick, 1961, 1962). The fact that *S*s did as well when the signal-to-interference ratio was 1 as under the no-interference condition and also showed no significant differences between the noise and blank-interval conditions might be due to the nature of the noise used. Using easily discriminable, highly overlearned symbols to represent both the signal and noise probably enhances *S*'s ability to differentiate and put the symbols into categories of signal and noise (Bruner, Goodnow, & Austin, 1956; Dollard & Miller, 1950). The simplicity of the task might also be a factor in the negative findings.

Since the results were negative for the effects of degree of noise, one would not expect any improvement due to practice, and the results confirm this. The lack of any significant effects due to the sequential interference of noise supports the hypothesis that a temporal separation of signal and noise does not effect signal detection, at least for the type of noise used in this study. Similar results were obtained in a study when a

signal and noise were simultaneously presented (Green et al., 1953); the color of a number to be searched for was specified in advance, and it was found that the presence of numbers of a clearly different color did not affect the search time. Archer (1954) used a visual task involving the identification of geometric patterns confounded with simultaneous noise and also found negative results in terms of the interference of the noise with the signal.

The predictions made on the basis of previous investigations by the author (Erlick, 1962) as to the amount of information that would be transmitted at the signal intensities of S_1 and S_2 were very accurate ($\pm .02$ bits). This lends validity to the proposed functional relationship between information transmitted in terms of identifying the more frequent category and the relative increment in frequency between categories. The fact that noise did not interfere with performance at either the S_1 or S_2 signal levels suggests that this relationship is constant for varying degrees of task difficulty.

Based on previous research by the author (Erlick, 1961), an improvement in perceiving the more frequent category of a random sequential series would be predicted as a function of decreasing the rate of presentation of the categories. It might be argued that the interspersing of blank intervals between the signal is equivalent to reducing the rate of presentation, i.e., a signal-to-interference ratio of 1 in this study reduces the average rate of signal presentation from 4 per sec. to 2 per sec. Consequently, an improvement in performance would be hypothesized as the signal-to-blank interval ratio increased. That no improvement in performance was found, might be explained in terms of the intervals being randomly interspersed throughout the signal instead of appearing as longer regular intervals as in the investigation of effects of rate. The lack of any significant differences between the noise conditions and the blank-interval conditions appears to indicate that for this task, the human could filter noise as effectively as the equipment.

Summary

The ability of human *Ss* to filter out noise from a task when the signal and noise are presented sequentially was investigated. Two levels of signal intensity and five signal-to-noise ratios (ranging from 1 to 22) plus a no-noise condition were used. A condition using blank time intervals instead of noise at all signal-to-noise ratios was also used. The task involved the judgment of the relative frequency of two random sequential categories (letters *A* and *B* presented visually). Ten different letters randomly interspersed among *A*s and *B*s represented the noise. There were no significant differences among any of the foregoing conditions, indicating that *Ss* effectively filtered out the noise when it was presented in sequential relationship to the signal.

References

Archer, E. J. Identification of visual patterns as a function of information load. *J. exp. Psychol.*, 1954, **48**, 313–317.

Archer, E. J., Bourne, L. E., Jr., & Brown, F. G. Concept identification as a function of irrelevant information and instructions. *J. exp. Psychol.*, 1955, **49**, 153–164.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. *A study of thinking.* New York: Wiley, 1956.

Dollard, J., & Miller, N. E. *Personality and psychotherapy.* New York: McGraw-Hill, 1950.

Erlick, D. E. Judgments of the relative frequency of a sequential series of two events. *J. exp. Psychol.*, 1961, **62**, 105–112.

Erlick, D. E. Perception of the most frequent category of a random series as a function of the number of categories. *J. exp. Psychol.*, 1962, **63**, 115–118.

Green, B. F., McGill, W. J., & Jenkins, H. M. The time required to search for numbers on large visual displays. *MIT Lincoln Lab. tech. Rep.*, 1953, No. 36.

Gregg, L. W. The effect of stimulus complexity on discriminative responses. *J. exp. Psychol.*, 1954, **48**, 289–297.

Hodge, M. H. The influence of irrelevant information upon complex visual discrimination. *J. exp. Psychol.*, 1959, **57**, 1–5.

Miller, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychol. Rev.*, 1956, **63**, 81–97.

Morin, R. E., Forrin, B., & Archer, W. Information processing behavior: The role of irrelevant stimulus information. *J. exp. Psychol.*, 1961, **61**, 89–96.

(Early publication received August 31, 1961)

Journal of Experimental Psychology
1962, Vol. 63, No. 2, 115–118

PERCEPTION OF THE MOST FREQUENT CATEGORY OF A RANDOM SERIES AS A FUNCTION OF THE NUMBER OF CATEGORIES[1]

DWIGHT E. ERLICK

Aerospace Medical Laboratory, Wright-Patterson Air Force Base

An earlier paper by the author (Erlick, 1961) presented findings which indicated that perception of the more frequent category of a sequence of two different categories is a function of the sequence of the categories, the duration of observation, and the rate of presentation. When both categories are equally frequent, the category having the higher degree of clustering is estimated as having a higher relative frequency of occurrence. Also, when both categories are presented randomly, and one occurs more frequently, the percentage of the correct identifications of the more frequent category increases as a function of an increasing difference in the relative frequencies between the categories. The ability to identify the more frequent category was found to decrease as a function of increasing rate of presentation (from one to eight events per sec.) and decreasing duration of observation (from 40 to 10 sec.). The foregoing investigation used only two categories. In terms of information analysis (Shannon & Weaver, 1949), the uncertainty as to which category had occurred more frequently was equivalent to 1 bit.

Many investigations, as reviewed by Miller (1956), have been concerned with the relation of amount of information input to information transmission. The findings generally indicate that as information input increases there is an initial increase in information transmission, then a leveling off at what is called channel capacity, and finally a decrease in transmitted information as overloading takes place.

The purpose of the present investigation was to increase the range of information inputs and determine two relationships: that between information transmitted and the frequency difference between categories; and that between information transmitted and information input.

METHOD

The same general procedure was used as in previous experiments (Erlick, 1961). A punched paper tape reader presented a preprogramed series of the mutually exclusive categories (symbols) on a single plane, in-line display unit. The unit contained 12 lenses, each etched with one of the numerical symbols. By turning on and off individual lights behind each of the lenses, the symbols were projected one at a time, so as to be centered in the same location of a 3×4 in. frosted plastic surface. The symbols, which were $3\frac{3}{4}$ in. high, were seen as an outline of light against a dark background. A motor-driven cam supplied pulses to the tape reader which controlled the rate of presentation of the symbols. Each symbol was presented for approximately half of each cycle, producing an equal on-off ratio. Based on the author's previous work a rate of four symbols per sec. was selected. This rate was well beyond *Ss* ability to count each symbol as it appeared, yet was sufficient so that a distinct off-time could be noticed when the same synbol was repeated in sequence. All of the programing equipment was housed in a separate room to reduce noise.

The number of different symbols presented

[1] The data were collected at the Engineering Psychology Research Project, Antioch College.

during one trial varied from 2 to 12 in four steps: viz., 2, 4, 8, or 12, equivalent to 1, 2, 3, and 3.6 bits of information input (H) respectively, and hereafter referred to as the H conditions. The numerical symbols were sets of consecutive integers beginning with 1, e.g., in the four-symbol condition, 1, 2, 3, and 4 were used. Within each condition a basic frequency of 20 was used for all except one symbol, which had an increased frequency of one of five values. Categories (symbols) having frequencies of 20 are designated by C_L, while C_M is used to represent the frequency of that category with the increment. The percentage increments between C_M and C_L are designated by ΔF in the formula:

$$\Delta F = \frac{C_M - C_L}{C_L} \times 100$$

Given a constant ΔF, it was hypothesized that as H increased, the accuracy of perceiving C_M would decrease. Therefore, in order to obtain somewhat equivalent accuracy in responses to the different H conditions, the ΔFs were increased as H increased (see Table 1).

TABLE 1

PERCENTAGE INCREMENTS (ΔF) OF THE MORE FREQUENT CATEGORY (C_M), PERCENTAGE OF CORRECT IDENTIFICATIONS OF C_M, AND INFORMATION TRANSMITTED (T), AS A FUNCTION OF INFORMATION INPUT (H)

H	ΔF	% Correct C_M Identifications	T
1	5	55.38	.01
	10	59.38	.03
	15	62.50	.05
	30	80.03	.28
	50	87.50	.46
2	10	30.90	.01
	20	35.42	.04
	30	50.87	.22
	50	64.93	.51
	90	82.99	1.07
3	15	20.66	.04
	30	32.64	.20
	50	37.85	.30
	100	63.89	1.04
	180	86.11	2.03
3.6	15	12.67	.02
	30	18.58	.08
	50	27.95	.24
	125	58.85	1.18
	210	79.69	2.15

A trial required one judgment and consisted of *S* observing a random series of either 2, 4, 8, or 12 different categories at one of the ΔF values. A session was restricted to the same H condition and consisted of three replications of each of five ΔFs randomly presented. An experimental period consisted of four sessions, each with one of the four different H conditions. The four H conditions were presented to four groups of 24 *S*s each. A latin square design was used to counterbalance order. For each H condition all the categories were counterbalanced over sessions and groups so that each occurred as the more frequent (C_M) category the same number of times. At the end of each trial, *S* was instructed to write the number which appeared most frequently. The *S*s were told to guess if they were not sure of the answer and to keep all previous answers covered with a card with which they were provided. All 96 *S*s had two experimental periods separated by at least 1 day, and made a total of 30 judgments of the four H conditions. Sessions were punched on one continuous tape with an automatic stop programed between trials. The *E*, who sat in the room with *S*s, called out the number of each trial before it was presented and started the trial by pressing a button. The *S*s were run in groups of up to 10 at one time, with each *S* seated in an individual booth facing the display and viewing the symbols from a distance of from 6 to 9 ft. The *S*s were given three practice trials before each session. An attempt was made to motivate *S*s by offering a monetary bonus as a function of the number of correct judgments.

RESULTS

The following analysis was based on the summated responses of all 96 *S*s. The percentages of correct identifications for each of the five frequency differences (ΔF), at each of the four levels of information input (H) were converted into information transmitted (T), as seen in Table 1. This measure (T) is derived by assuming that all categories within an H condition have equal likelihood of being selected as C_M. Thus, for example, in terms of information analysis (Shannon & Weaver, 1949), 75% identification of one of two equally likely events (H = 1) trans-

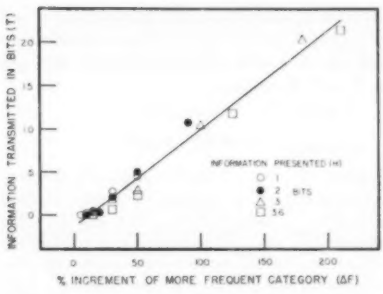

FIG. 1. Information transmitted as a function of the increment of the more frequent category for different information inputs. (Each point represents 576 observations.)

mits .19 bits, while 75% identification of one of eight equally likely events (H = 3) transmits 1.49 bits of information.

Graphic analysis indicated: no systematic interaction between symbol preference and the four H conditions; also no systematic effects for any of the experimental conditions from Session 1 to Session 2. Thus, the responses to trials in which different symbols appeared as C_M (for each ΔF within each H condition) were pooled over both sessions in the following analyses.

An examination of the relationship between T and ΔF (see Fig. 1) for each of the four H conditions of 1, 2, 3, and 3.6 bits gave the following equations based on the method of least squares with the subscript on T indicating H: $T_1 = .0107\ \Delta F - .0726$; $T_2 = .0138\ \Delta F - .1825$; $T_3 = .0123\ \Delta F - .2028$; $T_{3.6} = .0113\ \Delta F - .2427$. The overall equation based on all four H conditions is: $T = .0113\ \Delta F - .1342$, with the standard error of estimate, $\sigma_{T.\Delta F} = .09$.

Further analysis of the relationship between T and ΔF as indicated by the slopes of the equations shows approximately the same linear function for all H conditions. Using the slope of the overall equation, .0113, an increment of approximately 88% in frequency of a category is required to transmit 1 additional bit of information in terms of identifying a category as having occurred more frequently. Analysis of the relationship between T and H as indicated by the intercepts of the equations shows a slight decrease in T as H increases.

DISCUSSION

The lack of any systematic effects from Session 1 to Session 2 indicates no practice effects. The equations show approximately the same slope between T and ΔF for all four H conditions (1, 2, 3, and 3.6 bits) and there is a small standard error of estimate of the overall equation derived from all H conditions. Therefore, the ability to perceive the most frequently occurring category in a random series can be expressed as a linear relationship between reduction of uncertainty in identifying the more frequent category and the frequency differences between categories, at least for a range from 2 to 12 categories.

The lack of perfect information transmission even at the higher frequency increments is probably due in part to errors in recording by *Ss*. Current investigations by the author involving absolute estimation of proportions show a wide disparity between actual and estimated proportions for some of the judgments, e.g., an actual proportion of 85% is occasionally estimated as 10%.

The slight decrease in T as H increased indicates that an overload condition possibly exists at all H conditions. This is not surprising in view of the fast rates of presentation of the categories.

Using the slope (.0113) between T and ΔF based on all H conditions and the intercept for the H_1 condition, the result of a previous investigation (Erlick, 1951) involving the same variables but different *Ss* is predicted within $\pm .05T$. This finding adds validity to the foregoing relationships between T, ΔF, and H.

Summary

The ability to perceive the more frequently occurring category in a random series as a function of the number of different categories was investigated. The main variables of concern were: the relationship between information transmitted (T) and (*a*) the frequency difference between categories (ΔF), and (*b*) the information input (H). The different categories were 2, 4, 8, or 12 (H in bits = 1, 2, 3, and 3.6, respectively). Within each H condition a basic frequency of 20 was used for all except one category, which had an increased frequency of one of five values. Categories were presented randomly at the rate of four per sec. The *S*s were instructed to indicate the category which appeared most frequently. The results indicate that approximately the same linear relationship exists between T and ΔF for all H conditions with T = .0113 ΔF − .1342. The relationship between T and H indicates a slight decrease in T as H increases.

References

Erlick, D. E. Judgments of the relative frequency of a sequential series of two events. *J. exp. Psychol.*, 1961, **62,** 105–112.

Miller, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychol. Rev.*, 1956, **63,** 81–97.

Shannon, C. E., & Weaver, W. *The mathematical theory of communication.* Urbana: Univer. Illinois Press, 1949.

(Early publication received August 14, 1961)

Journal of Experimental Psychology
1962, Vol. 63, No. 2, 119-123

MANIFEST ANXIETY SCALE SCORE AND THE READY SIGNAL IN CLASSICAL CONDITIONING[1]

WILLIAM F. PROKASY, JR. AND FRANCIS L. WHALEY

Pennsylvania State University

With the construction of the Taylor Manifest Anxiety scale (Taylor, 1953), a number of studies (Baron & Connor, 1960; Caldwell & Cromwell, 1959; Spence & Farber, 1953; Spence, Farber, & Taylor, 1954; Spence & Taylor, 1951, 1953; Spence & Weyant 1960; Taylor, 1951) have reported a positive relationship between conditionability and score on the *MA* scale (Manifest Anxiety scale). Similarly, Spence and Beecroft (1954) and Spence and Farber (1954) report greater responsiveness to the reinforced stimulus in the classical discrimination learning situation, as well as some trend (nonsignificant) for better discrimination between reinforced and nonreinforced stimuli, in *Ss* who scored high, rather than low, on the *MA* scale.

A problem is posed, however, by the studies reporting negative, or reverse findings. Hilgard, Jones, and Kaplan (1951) found greater classical discrimination differences in low-*MA* *Ss*, rather than in high. Prokasy and Truax (1959) reported no difference between high- and low-*MA* score groups, while King (1958) found low-*MA* score *Ss* averaging significantly greater responding than did high-*MA* score *Ss*.

It is the King (1958) study which provides the clue for an explanation of these discrepancies. In a standard eyelid conditioning situation, half of the *Ss* were high-*MA* scorers and the other half were low-*MA* scorers. Each of these groups was subdivided into two subgroups: one subgroup received a "ready, blink" signal 2 to 4 sec. before each conditioning trial, and the other subgroup did not receive this warning signal. In addition to the surprising (and to date unexplained) observation that in the presence of the ready signal the low-*MA* *Ss* were the more frequent responders, it was found that there were no differences between high- and low-*MA* *Ss* without the ready signal preceding each trial.

Thus, studies using a ready, blink signal have uniformly found that high-*MA* *Ss* respond more frequently and tend to discriminate better than low-*MA* *Ss*. When a blink was not required (Hilgard et al., 1951), the latter condition did not obtain. When a ready signal was not employed, the results are conflicting: Baron and Connor (1960) and Spence and Weyant (1960) report positive results, while Prokasy and Truax (1959) and King (1958) report negative results.

It is the purpose of this study to explore further the role of the ready, blink signal as a potential interacting factor in the high- vs. low-*MA* score conditioning studies reported earlier, with present emphasis placed upon classical discrimination learning.

METHOD

Apparatus.—The eyeblink conditioning apparatus has been described in detail else-

[1] This study was supported by funds from National Institute Mental Health Grant M-2560A and from National Science Foundation Grant 7463 to the senior author. The authors wish to acknowledge the technical assistance provided by Burton Klinger and Patricia Fitzgerald.

where (Prokasy & Whaley, 1961). The *Ss* were seated from 30 to 40 in. in front of a panel including a ½-in. orange fixation light (445 ml.) mounted midway between two milkwhite disks of 3-in. diameters. The disks were mounted vertically, their centers being separated by 8 in. The reinforced and unreinforced CSs were increases from 3.8 to 190 ml. illumination of the upper or lower disk. For half of the *Ss* the upper disk was reinforced, and for half the lower disk was reinforced. The UCS was 1.5 liter/min puff of nitrogen presented through a 5-mm. (inside diameter) glass jet situated within .25 in. of *S*'s eyelash. Trials were automatized by a Western Union teletype transmitter with a fixed intertrial interval of 15 sec. Hunter-Brown interval timers controlled the interstimulus interval of 500 msec., the CS duration of 500 msec., and the UCS duration of 50 msec. The CS and UCS signals, the response, and the first time derivative of the response were recorded by a capillary pen recording system on a paper drive operated at 150 mm. per sec.

Subjects.—At the beginning of the semester all students in two large introductory classes at Pennsylvania State University were given the *MA* scale inventory under the guise of a "reliability" study. Those students scoring at and below 9 or at and above 21 were later requested to serve as *Ss* in an experiment, but were told that they had been selected from the class rolls "at random." The means and *SDs*, respectively, of the two *MA* score groups are: High-*MA* score group, 26.19, 4.63; Low-*MA* score group, 6.7, 1.77. A total of 160 *Ss* participated in this experiment, and, of these, 8 were lost because of apparatus failure or *E* error.

Procedure.—The two independent variables are the presence or absence of a ready, blink signal before each trial and two ranges of *MA* scores. As *Ss* came to the laboratory they were assigned alternately to a signal or a no-signal condition. To avoid any bias in the situation, *E* did not have, at the time of running *Ss*, information on whether they fell into the high- or low-*MA* score groups. As a result, the numbers of *Ss* in the signal and no-signal partitions were equal, but they were not equal within *MA* conditions. Thus, there were four groups of *Ss*: High *MA*, ready signal with 44 *Ss*; Low *MA*, ready signal with 32 *Ss*; High *MA*, no ready signal with 38 *Ss*; and Low *MA*, no ready signal with 38 *Ss*. Of the 152 *Ss* remaining, 15 met the criterion defined by Spence and Taylor (1951) as "voluntary responders" (i.e., half or more of their CRs fell in the latency range 200 to 300 msec.), and were not employed in the analyses. Seven voluntary responders were in the no-signal groups (5 high *MA* and 2 low *MA*) and 8 were in the signal groups (3 high *MA* and 5 low *MA*). In order to restore cell proportionality for statistical analyses, the data of 4 *Ss* from the high-*MA*, ready-signal group and 3 *Ss* from the low-*MA*, no-ready-signal group were randomly eliminated. This left a total of 60 *Ss* in the low *MA* partition (30 ready signal and 30 no ready signal) and 70 in the high *MA* partition (35 ready signal and 35 no ready signal). All records were scored by a paid undergraduate assistant who did not know the nature of the study.

The *Ss* were given a total of 90 trials, 45 to the reinforced CS and 45 to the nonreinforced CS, with positive and negative stimuli presented in an unsystematic order. Those *Ss* in the ready, blink signal condition received this signal from 2 to 4 sec. prior to CS onset. The *Ss* were instructed to keep the fixation light in view at all times and to maintain a passive attitude. After the session each *S* was cautioned not to discuss the experiment with other students, although

TABLE 1

MEANS AND *SDs* OF RESPONSE FREQUENCY TO LAST 20 REINFORCED AND LAST 20 NONREINFORCED STIMULI

MA	Signal				No Signal			
	Reinforced		Nonreinforced		Reinforced		Nonreinforced	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
High	10.89	5.71	7.12	4.82	12.03	4.06	8.89	4.54
Low	7.90	4.87	4.50	4.11	11.30	4.38	8.13	4.15

none was aware of the true purpose of the experiment.

Results

Mean numbers of responses to the reinforced and nonreinforced stimuli over the last 40 trials (20 reinforced and 20 nonreinforced) for all four groups are presented in Table 1. Employing the difference between numbers of responses to the reinforced and nonreinforced stimuli as the index of discrimination learning, an analysis of variance was performed on the discrimination learning data. Based upon an average within groups mean square of 22.34, F ratios as a function of the MA and signal variables, as well as their interaction, were all less than unity. A post hoc analysis based upon 36 Ss who scored above 26 and 24 Ss who scored below 6 on the MA scale resulted in no additional differentiation among treatment combinations.

Table 2 provides a summary of an analysis of variance of total number of responses to the reinforced stimulus on the last 20 reinforced trials. Overall responding in the high-MA group is significantly higher than in the low-MA group, and overall responding in the no-signal group is significantly greater than in the signal group. As suggested by the means in Table 1, high- and low-MA groups did not significantly differ in the no-signal condition ($F = .52$; $df = 1/63$) but did significantly differ in the signal condition ($F = 5.1$; $df = 1/63$).

TABLE 2

Analysis of Variance of the Number of Responses on the Last 20 Reinforced Trials

Source	*df*	*MS*	*F*
MA level	1	111.43	4.82*
Ready signal	1	155.11	6.71*
MA × Signal	1	41.15	1.78
Within groups	126	23.12	—

* $P < .05$.

Discussion

The findings indicate that prior results with the MS variable may be dependent upon anticipatory effects induced by the ready signal. The data of three prior studies, however, are not consistent with this conclusion. The first is that of Hilgard, Jones, and Kaplan (1951), in which the high-MA score group manifested less discrimination between reinforced and nonreinforced stimuli than did the low-MA score group. While these investigators employed the ready signal, they did not require that S blink before CS onset as is customary in the studies from the Iowa laboratory and as was done in the present study. The possibility that a blink response between signal and CS onset interacts with the MA score conditionability relationship remains to be explored.

The second such study is that of Spence and Weyant (1960) who compared the conditioning performances of high- and low-MA score groups without preceding each trial with a ready signal. In accord with past studies from the Iowa laboratory (cited in the introduction), the high-MA score group manifested a higher response frequency. Since their finding without the ready signal differs from those of King (1958), Prokasy and Truax (1959), and the present study, it is, perhaps, worth pointing out that Spence and Weyant introduced an irrelevant discrimination task into the conditioning situation in order to mask the nature of the experiment. Between most conditioning trials, Ss were required to press a telegraph key in response to an auditory stimulus. Here, again, the possibility of an interacting influence cannot be overlooked, as the effects upon conditioning of an irrelevant discrimination task are unknown.

Finally, Baron and Connor (1960) obtained a rho of .4 between response frequency and *MA* score in classical eyeblink conditioning, but did not employ a ready signal preceding each trial. These results are similar to those of Taylor (1951) and it is not clear why they differ so markedly from those of King (1958) and the present investigators.

In conjunction with the present research, the Baron and Connor (1960) study suggests that the *MA* scale may separate *Ss* in two ways: distinguishing high from low conditioners and, in addition, separating those *Ss* who are made more susceptible to conditioning by virtue of a verbally induced preparatory set. Whether or not the effects of either aspect of the *MA* scale becomes manifest, however, may well depend upon rather subtle influences in the conditioning situation: influences such as specific UCS intensity, interstimulus interval, intertrial interval, or, possibly *E-S* rapport.[2]

That overall depression of response frequency was obtained in the presence of the ready signal is consistent with the results of studies by King (1958) and by McAllister and McAllister (1960), but is not compatible with the position of Spence and Weyant (1960) who posit at least a small increment in drive to be initiated with the presentation of the ready signal. Such a drive increment would, at the very least, lead to equal or greater performance with the ready signal than without it.

In conclusion, the observations that the presence of the ready signal reduces the overall level of responding and interacts with the manifestation of *MA* related conditionability suggests that the ready signal be more fully explored before it is adopted as a standard procedure in the classical conditioning situation. While it is true that there is no one "real" conditioning procedure, it appears evident that the ready signal is not a necessity for the development of CRs. To add it as part of conditioning methodology would seem to require that the investigator have a reasonable understanding of its operation as an interacting independent variable.

[2] G. A. Kimble (personal communication) and M. R. Baron (personal communication) have indicated that *E-S* interaction may play a role in the *MA* conditionability relationship.

Summary

With high and low *MA* scores and presence or absence of a ready signal as the independent variables, 152 *Ss* served in a classical discrimination learning situation. The principal findings were (*a*) that the ready signal depressed overall frequency of responding, and (*b*) that high-*MA* score *Ss* responded significantly more than did low-*MA* score *Ss* only under the signal condition.

References

Baron, M. R., Connor, J. P. Eyelid conditioned responses with various levels of anxiety. *J. exp. Psychol.*, 1960, **60**, 310–313.

Caldwell, D. F., Cromwell, R. L. The relationship of manifest anxiety and electric shock to eyelid conditioning. *J. exp. Psychol.*, 1959, **57**, 348–349.

Hilgard, E. R., Jones, L. V., & Kaplan, S. J. Conditioned discrimination as related to anxiety. *J. exp. Psychol.*, 1951, **42**, 94–99.

King, M. S. Anxiety and eyelid conditioning: The effect of a ready signal. Unpublished master's thesis, Duke University, 1958.

McAllister, W. R., & McAllister, D. E. The influence of the ready signal and unpaired UCS presentations on eyelid conditioning. *J. exp. Psychol.*, 1960, **60**, 30–35.

Prokasy, W. F., & Truax, C. B. Reflex and conditioned responses as a function of manifest anxiety. *Amer. J. Psychol.*, 1959, **72**, 262–264.

Prokasy, W. F., & Whaley, F. L. The intertrial interval variable in classical conditioning. *J. exp. Psychol.*, 1961, **62**, 560–564.

Spence, K. W., & Beecroft, R. S. Differential conditioning and level of anxiety. *J. exp. Psychol.*, 1954, **48**, 399–403.

Spence, K. W., & Farber, I. E. Conditioning and extinction as a function of anxiety. *J. exp. Psychol.*, 1953, 116–119.

SPENCE, K. W., & FARBER, I. E. The relation of anxiety to differential eyelid conditioning. *J. exp. Psychol.*, 1954, **47**, 127–134.

SPENCE, K. W., FARBER, I. E., & TAYLOR, E. The relation of electric shock and anxiety to level of performance in eyelid conditioning. *J. exp. Psychol.*, 1954, **48**, 404–408.

SPENCE, K. W., & TAYLOR, J. A. Anxiety and strength of UCS as determinants of amount of eyelid conditioning. *J. exp. Psychol.*, 1951, **42**, 183–188.

SPENCE, K. W., & TAYLOR, J. A. The relation of conditioned response strength to anxiety in normal, neurotic, and psychotic subjects. *J. exp. Psychol.*, 1953, **45**, 265–272.

SPENCE, K. W., & WEYANT, R. G. Conditioning performance of high- and low-anxious *Ss* in the absence of a warning signal. *J. exp. Psychol.*, 1960, **60**, 146–149.

TAYLOR, J. A. The relationship of anxiety to the conditioned eyelid response. *J. exp. Psychol.*, 1951, **41**, 81–92.

TAYLOR, J. A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, **48**, 285–290.

(Received October 28, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 2, 124–128

A TEST OF TWO INTERPRETATIONS OF THE APPARENT SIZE EFFECTS IN A DISTORTED ROOM [1]

WILLIAM EPSTEIN

University of Kansas

If two objects of equal physical size are placed in the two rear windows of a monocular distorted room, they will appear to be unequal in size. One explanation of such size effects which is frequently cited is formulated in terms of an interaction of assumptions, apparent distance, and retinal size. This interpretation is often associated with the transactionalist school (Kilpatrick, 1952) and may be called the *neoempiristic* viewpoint.

An alternative explanation is possible which makes no mention of assumptions or apparent distance. Instead, the effects are seen to depend entirely on the relational properties of the proximal stimulus configuration. This interpretation may be designated as the *relational* viewpoint.[2] If apparent size is relationally determined, then it follows that when the relational stimulus patterns which are being compared are not identical, the perceived size will not be identical as well. In the distorted room the two rear windows are retinally equal because the size of the farther window has been increased in proportion to the added distance. However, the retinal sizes of the physically identical objects whose size is to be judged are not equal. The more distant object projects a much smaller image. Therefore, the proximal stimulus relationships in the two fields are different, and the object in the more distant window will appear to be smaller. This outcome is determined by the retinal size of the stimulus objects relative to the retinal size of their respective frameworks. If this is indeed the case, then the introduction of the factors of apparent distance and "assumed normalcy" becomes superfluous.

The purpose of the present investigations was twofold: (*a*) a quantitative estimate of the magnitude of the anomalous effects was sought, and (*b*) an attempt was made to determine whether relational factors would account for the apparent size effects.

Experiment I

The literature dealing with the distorted room (e.g., Ittelson & Kilpatrick, 1952; Kilpatrick, 1954; Wiener, 1956; Wittereich, 1959) reveals a common tendency to describe the obtained size distortions qualitatively. For this reason, Exp. I was designed to provide a quantitative estimate of the distortions of apparent size.

Method

Apparatus.—The main apparatus is pictured and described in detail elsewhere (Ittelson, 1952, pp. 40–43; Ittelson & Kil-

[1] The work described in this paper was supported by a grant (M-4153) from the National Institute of Mental Health of the United States Public Health Service and a grant from the General Research fund of the University of Kansas. Thanks are due to Raymond Engstrand who collected the data and assisted in the statistical computations.

[2] See Rock and Ebenholtz (1959) for a recent discussion of the relational determination of apparent size.

patrick, 1952, pp. 41–55). It was the distorted room labeled "Monocular Distorted Room No. 1" in Ittelson's (1952, p. 41) manual. For the present purposes the following details are necessary. The back wall was tilted from the frontal parallel plane from right to left away from *S*. The distance from the reference point to the center of the base of the right window was 40.25 in., while the distance to the same point in the left window was 46.38 in. The mean heights of the right and left windows were 19.50 in. and 26.50 in., respectively.

Two rods, painted white, were placed in position at the objective centers of the bases of the two windows. The rod in the right window was 0.5 in. in diameter, while the rod in the left window was 0.625 in. in diameter. During each judgment the rod which served as the standard was kept at a constant height of 10 or 15 in. By turning a wheel *S* could produce changes in the height of the comparison rod.[2] At one extreme the rod could be eliminated from view, and at the other extreme it could be made to span the entire length of the window.

Subjects.—The *S*s in Exp. I were 40 students of introductory psychology. None of the *S*s wore corrective glasses and all indicated that they had normal vision.

Procedure.—Each *S* was assigned to one of the following four experimental conditions: Cond. 1a ($N = 10$): Standards on left, 15-in. standard first; Cond. 1b ($N = 10$): Standards on left, 10-in. standard first; Cond. 2a ($N = 10$): Standards on right, 15-in. standard first; and Cond. 2b ($N = 10$): Standards on right, 10-in. standard first.

Two ascending and two descending judgments of apparent equality for each standard were obtained by the equation method. The judgments for each standard were made successively in a balanced order. The comparison stimulus was set at an initial height of 2 in. from the bottom of the window for the ascending trials; and for the descending trials it was set at a height which was 2 in. from the top of the window. No time limit was placed on *S*'s setting, but a record was kept of the time required for each setting. The times ranged from 5 to 10 sec.

When *S*'s head was in the desired position he was asked to describe the room. Neither of the two rods was visible during this period.

[2] The measuring device was not visible to *S*, nor was there any opportunity for associating any specific position of the wheel with a particular size.

TABLE 1

MEAN JUDGMENTS (IN.) OF APPARENT EQUALITY IN EXP. I AND II

Size of Standard	Condition 1		Condition 2	
	Mean	*SD*	Mean	*SD*
10 in.	7.31	0.24	13.74	0.29
15 in.	10.93	0.33	20.73	0.27
8 in.	6.22	0.56	10.22	1.28
12 in.	9.26	0.72	15.21	1.81

Then the instructions were read by *E*. The instructions encouraged *S* to base his judgments on immediate, uncalculated perceptual impressions of size.

After the final judgment *S* was questioned about the appearance of the distorted room and also about the procedure for arriving at a satisfactory match.

Results

The mean judgments of apparent equality were nearly identical for Cond. 1a and 1b, and the same was true for Cond. 2a and 2b. For this reason the results for Cond. 1a and 1b have been combined, and the same has been done for Cond. 2a and 2b. A summary of the data is contained in the upper two rows of Table 1. Although these results were not unexpected, they are nonetheless quite striking. If apparent equality can be considered to be a transitive relationship, then the data for the 10-in. and 15-in. standards, respectively, may be expressed in the following way: 7.31 in. = 10.00 in. = 13.74 in., and 10.93 in. = 15.00 in. = 20.73 in.

All of the distributions were highly homogeneous as indicated by the *SD*s given in Table 1. Similar homogeneity was observed in *S*s' descriptions of the distorted room. All *S*s described it as a room whose appearance was completely normal. The postexperimental interview revealed that *S*s were unaware of the objective nature of the stimulus situation, as

they were at the outset of the experiment.

Experiment II

In order to determine whether the size effects were relationally determined, a stimulus situation was created in which the neoempiristic and relational viewpoints would predict opposed, mutually exclusive outcomes.

Method

Apparatus.—The apparatus was the same as in Exp. I with the exception of three modifications: (*a*) The major change consisted of reducing the size of the left, more distant window to the dimensions of the right, nearer window. (*b*) Both rods were 0.5 in. in diameter and were painted the same medium gray as the window frames. (*c*) The standard rods were 8 and 12 in. in height.

Subjects.—The *S*s were 48 students from the same population. All had normal, uncorrected vision.

Procedure.—The procedure was identical to that followed in Exp. I with the exception of two minor alterations. The following sentence was added to the instructions: "It is equally important that when you have completed your match the rods look to you to be equal in length with respect to each other, and not necessarily with respect to the windows."[4] The second difference in procedure was an addition to the post-experimental interview. The *S* was requested to "draw the room as you remember seeing it."

Under the modified stimulus conditions of this experiment the two interpretations require different outcomes. Assuming that the normal rectangular appearance of the room is unaffected by the alterations described above, then the neoempiristic theory predicts size matches comparable to those obtained in Exp. I. Relational considerations lead to a different prediction. The standard and comparison should appear to be equal when they are in fact objectively equal. This requirement follows from the observation that only when the two rods are objectively equal will the two retinal fields be identical relationally. The same considerations apply irrespective of the location of the standard and, therefore, the prediction holds both for Cond. 1 and Cond. 2.

Results

Eight *S*s showed by their pre-experimental comments and post-experimental drawings that they had perceived the true shape of the distorted room. Therefore, the data reported for this experiment are based on the responses of only 40 *S*s who were evenly distributed among the four experimental conditions. None of the 40 *S*s whose responses were accepted reported any perceived distortion. There was only one major difference in the appearance of the room between the two experiments. In Exp. I both windows were perceived as equal in size; in Exp. II the left window appeared to be considerably smaller than the right window.

The results for Cond. 1a and 1b have been combined, and the results of Cond. 2a and 2b have been combined. A summary of these data is contained in the bottom two rows of Table 1. The results were in close accord with the neoempiristic expections and are comparable to those obtained in Exp. I.

Discussion

The outcome of Exp. I is compatible with both interpretations. However, the results of Exp. II are consonant only with the neoempiristic viewpoint. The neoempiristic thesis requires that the adjustments in the physical size of the comparison should be in direct proportion to the difference in radial physical distance between the rods. The requirements of size-distance proportionality for both experiments are contained in Table 2 together with the deviation values, i.e., the amount of discrepancy between the obtained (see Table 1)

[4] This segment of the instructions was adopted with only minor changes from Rock and Ebenholtz (1959, pp. 389–390). The original version refers to "lines" and "frames" instead of rods and windows.

TABLE 2

VALUES (IN.) REQUIRED BY NEOEMPIRISTIC INTERPRETATIONS AND PERCENTAGE DEVIATIONS FROM THE REQUIRED VALUES

Size of Standard	Condition 1		Condition 2	
	Required Value	% Deviation	Required Value	% Deviation
10 in.	8.50	−14	11.50	+19
15 in.	13.50	−19	17.50	+18
8 in.	7.00	−11	9.00	+14
12 in.	10.50	− 7	14.00	+ 8

Note.—% Deviation is $\frac{\text{Req. Value}}{\text{Deviation}}$. A negative sign indicates that the obtained match was short of the required height. A positive sign indicates that the match exceeded the required height.

and the required values. In Exp. I the magnitude of the deviations from the values required by the neoempiristic view ranged from −14% to +19%; in Exp. II the range was −11% to +14%.

There are several possible reasons for the ineffectivess of relational factors in the present situation: (*a*) Unlike the Rock and Ebenholtz (1959) studies, the rods and their immediate frameworks were not presented in isolation. Both stimulus fields were perceived simultaneously within a common frame of reference, i.e., the back wall. The simultaneous presence of several subordinate and superordinate frameworks may have vitiated the influence of the immediate framework. (*b*) Another reason concerns the logical organization of the experimental task. Under Cond. 1, *S* is shown a standard which spans a specified length of a phenomenally small window. On his right *S* sees a window which is subjectively much larger. The relational view requires that the comparison in the right window be made objectively equal to the standard to achieve apparent equality. In fact, the ideal *S* would adjust the size of the comparison rod so that it spanned the same proportion of the window as does the standard in its window. However, a setting of this nature means that two rods will be judged equal in size when they span equal proportions of *phenomenally differently sized windows.* This is not a logically tenable outcome. Nevertheless, it is the result predicted by the relational point of view. Thus, the relational effect is opposed by the necessity for internal consistency and, as a result, may have been overriden. However, it should be noted that the instructions in Exp. II were deliberately formulated in a manner which was intended to discourage intentional judgmental manipulation. The *S*s were asked to make an immediate judgment of apparent size, and very few indicated that they had done otherwise. The fact that all *S*s completed their judgments in 10 sec. or less is presumptive evidence against the intrusion of deliberate judgment.

The present investigation gives little reason to question the critical role of apparent distance in producing the anomalous size effects obtained in the distorted room. However, there is one reservation which should be mentioned with regard to the usual formulation of the neoempiristic thesis. It is hard to see the justification or necessity for introducing "assumptions," "predicdictions," or "involuntary bets" into the explanation. The results seem to be explicable entirely in terms of an interaction between retinal size and apparent distance.[5]

SUMMARY

The anomalous apparent size effects observed in a distorted room may be explained in two ways. The *neoempiristic* interpretation explains the effects in terms of an interaction of assumptions, apparent distance, and retinal size. The relationalist interpretation refers the effects to the relational properties of the proximal situation.

Two experiments concerned with the anomalous apparent size effects in a monocular distorted room were reported. Experiment I provided a quantitative estimate of the obtained size distortions. The data were compatible with both the neoempiristic and the relationalist viewpoints.

[5] See Linksz (1952, pp. 628–629) for a complete exposition of this point.

In Exp. II a stimulus situation was created which made possible a test of the two explanatory alternatives. The construction of the distorted room was modified in a manner as to lead the two interpretations to predict mutually opposed size distortions. The results fit the neoempiristic interpretations but were not consistent with an explanation in relational terms.

References

Ittelson, W. H. *The Ames demonstrations in perception.* Princeton: Princeton Univer. Press, 1952.

Ittelson, W. H., & Kilpatrick, F. P. Equivalent configurations and the monocular and binocular distorted rooms. In F. P. Kilpatrick (Ed.), *Human behavior from the transactional point of view.* Hanover, N. H.: Institute of Advanced Studies, 1952.

Kilpatrick, F. P. (Ed.) *Human behavior from the transactional point of view.* Hanover, N. H.: Institute of Advanced Studies, 1952.

Kilpatrick, F. P. Two processes in perceptual learning. *J. exp. Psychol.*, 1954, **47**, 362–270.

Linksz, A. *Physiology of the eye.* Vol. 2. *Vision.* New York: Grune & Stratton, 1952.

Rock, I., & Ebenholtz, S. The relational determination of perceived size. *Psychol. Rev.*, 1959, **66**, 387–401.

Wiener, M. Perceptual development in a distorted room: A phenomenological study. *Psychol. Monogr.*, 1956, **70**(16, Whole No. 423).

Wittereich, W. J. Visual perception and personality. *Scient. American*, 1959, **200**, 56–60.

(Received November 9, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 2, 129–142

CONFUSION MATRIX ANALYSIS OF SINGLE AND MULTIDIMENSIONAL AUDITORY DISPLAYS[1]

MILTON H. HODGE

University of Georgia

AND IRWIN POLLACK

Operational Applications Office, Bedford, Massachusetts

A major goal in psychophysical experimentation is the prediction of the complete pattern of confusions among alternative stimulus objects and their associated responses. An important step in this direction, the *constant-ratio rule* (CRR), has been furnished by Clarke (1957). The rule has been successfully applied in a number of situations employing "complex" informational displays, e.g., words or syllables presented in noise (Clarke, 1957; Clarke & Anderson, 1957; Pollack & Decker, 1960) and words flashed briefly with a tachistoscope (Anderson, 1959). However, a preliminary experiment (Clarke, 1959) with elementary auditory displays, e.g., the frequency and intensity of a tone, has suggested that the rule may break down. Considering the subtle complexities associated with the responses to such displays, the failure of the CRR—or any other single rule—is reasonable. Consequently, the present experiments were undertaken to provide a better foundation for making a decision about the rule with respect to elementary displays.

Definitions

Before examining the rule and some of its implications, consider the following definitions:

Stimulus ensemble.—A well-defined set of stimulus objects; e.g., eight sound levels of a 1000-cps tone, 0.25 sec. in duration, equally spaced on a db. scale between sound levels of 81 and 88 db.

Response set.—A well-defined set of responses corresponding one to one with the objects of the stimulus ensemble; e.g., the cardinal numerals 1–8 corresponding with the ensemble of eight stimulus objects.

Stimulus subensemble.—An arbitrary selection of stimuli from a given stimulus ensemble; e.g., any four or any two of the eight sound levels in the stimulus ensemble.

Single-dimensional ensemble.—A stimulus ensemble in which the stimulus objects differ only with respect to a single variable; e.g., the intensity or the duration of a tone.

Multidimensional ensemble.—A stimulus ensemble in which the stimulus objects differ with respect to more than a single variable; e.g., the frequency, intensity, and duration of a tone. Another example is speech.

Identification.—An experimental task in which *S* attempts to match members of the response set with corresponding members of the stimulus ensemble; the task required in the method of absolute judgment.

Stimulus-response confusion matrix.—A description of an identification experiment in terms of the entire distribution of responses to each member of the stimulus ensemble; e.g., the matrix of Table 1 is a 4×4 confusion matrix.

Correct response.—An identification response which matches the corresponding member of the stimulus ensemble; a response on the negative diagonal of the confusion matrix; e.g., the bold face entries of Table 1.

Incorrect response.—A response which does not match the corresponding member of the stimulus ensemble; a response not on the negative diagonal.

Observed confusability.—The mean proportion of observed incorrect responses associated with a designated stimulus subensemble.

Derived confusability.—The mean proportion of calculated incorrect responses associated with a designated subensemble after the arbitrary removal of the other members of

[1] This is Technical Report 60-43 of the Air Force Command Control Division. The research is in support of Project 7682, Information Processing by the Human Operator.

TABLE 1

A STIMULUS-RESPONSE CONFUSION MATRIX

Stimuli	Responses			
	1	2	3	4
I	*a*	*b*	*c*	*d*
II	*e*	*f*	*g*	*h*
III	*i*	*j*	*k*	*l*
IV	*m*	*n*	*o*	*p*

Note.—Derived confusability scores:

$$(\text{I, II}) = 0.5\left[\frac{b}{(a+b)} + \frac{e}{(e+f)}\right]$$

$$= \frac{b+e}{(a+b+e+f)} \quad \text{when } a+b = e+f$$

$$(\text{I, IV}) = 0.5\left[\frac{d}{(a+d)} + \frac{m}{(m+p)}\right]$$

$$= \frac{d+m}{(a+d+m+p)} \quad \text{when } a+d = m+p.$$

Each alphabetic entry represents the conditional probability of a specific response (indicated by the column number), given the stimulus (indicated by the row number). When stimuli are equally likely, the conditional response probabilities are proportional to the original matrix entry.

the ensemble and their associated responses; e.g., algebraically defined in Table 1 for two particular pairs of stimulus objects.

Contextual effects.—Changes in the derived confusability of a given subensemble which result from a particular definition and selection of other members of the stimulus ensemble; e.g., changes in derived confusability occurring as the result of changes in the spacing among the stimulus objects.

Constant-ratio rule.—An empirical rule which states that the derived and observed confusability of any particular subensemble are identical and that both are independent of the other members of the stimulus ensemble; e.g., if a specific subensemble were embedded in ensembles of different size, the rule would predict that the derived confusability would be invariant with the size of the ensemble.

Implications of the Constant-Ratio Rule

The CRR implies the lack of stimulus and response interactions other than those between pairs of stimuli and their associated responses. The confusability of a given stimulus pair and its associated responses is assumed to be independent of other stimuli and responses in the matrix. That is, when stimuli and responses are eliminated from the confusion matrix, the CRR assumes that the conditional probabilities of the remaining members of the matrix do not change. For example, consider the ensemble of four stimuli I, II, III, and IV with associated Responses 1, 2, 3, and 4, illustrated in Table 1. The CRR suggests the following operations in obtaining the derived confusability between I and II and their responses. We first obtain the initial 4×4 matrix as shown in Table 1. We then dispose of Stimuli III and IV and of Responses 3 and 4. Finally, the remaining 2×2 matrix is normalized so that its conditional response probabilities sum to 1.0. The rule predicts that the resulting derived confusability will be identical with the observed confusability of a matrix consisting of Stimuli I and II and Responses 1 and 2.

The postulated lack of stimulus and response interactions is in direct opposition to numerous psychophysical studies demonstrating adaptation-level, anchoring, and response-bias effects with single dimensional stimuli. These effects all imply that the interactions are a function of the stimulus ensemble and response set in which the stimuli and responses are embedded. An extreme example of the interaction approach is the rubber-band model (RBM) of sensory organization (McGill, 1960). In a particular discrimination situation, e.g., identification of the objects of an ordered stimulus ensemble, the model assumes that *S* has a given discriminative capacity available for the discrimination task. If confusability is employed as a measure of *S*'s performance of the task, the RBM states that confusability is inversely proportional to discriminative capacity. Other *S*s performing the

same task may have different confusability scores because of different discriminative capacities. The model further assumes that the discriminative capacity available for a given task is stretched over the range of the stimulus ensemble under examination. Schematically, imagine a rubber band with a fixed number of molecules (discriminative capacity), and capable of being stretched over a finite distance (the range of a stimulus ensemble). When stretched over a wide range, a lesser amount of the available discriminative capacity is assignable to a given stimulus separation (stimulus pair) than when stretched over a narrow range. Consequently, the confusability of two stimulus objects will increase with increases in the range of the ensemble. To summarize, the RBM assumes that the confusability of two stimulus objects is a function of the available discriminative capacity and the range of the stimulus ensemble. Confusability will increase with a decrease in discriminative capacity or with an increase in the ensemble range over which a given discriminative capacity must be stretched.

The CRR and the RBM thus represent opposite extremes of a continuum of postulated interaction, with the CRR assuming no interaction and the RBM assuming complete dependence upon the order and range of the stimulus ensemble. Although the assumptions of the RBM tend to restrict the model to single-dimensional ensembles, recent work with scaling problems (Torgerson, 1958) may make its application to multidimensional ensembles feasible. The CRR is not bound by the order and range assumptions, and is presumably applicable to both single and multidimensional ensembles.

The present experiments were designed to examine the accuracy of the CRR in predicting the performance of *S*s to a variety of auditory stimulus ensembles. More specifically, the experiments (*a*) evaluate contextual effects in single-dimensional ensembles, and (*b*) assess the adequacy of the CRR in both single and multidimensional ensembles.

General Method

In all, nine experiments were completed. The experiments are numbered in order of their execution, but for purposes of presentation, the separate experiments are reorganized.[2] Features common to the separate studies include:

Subjects.—A single crew of 7 university students served in all tests. They were tested in groups of 6, with each *S* enjoying a brief rest between successive experimental runs. All *S*s had extensive previous experience with psychophysical testing. The *S*s, reimbursed for their services, were employed 8 hr. a day over a period of 3 months. The first 10 testing days were devoted to preliminary practice in which asymptotic levels of performance were achieved.

Apparatus and procedure.—A stepping-switch programer selected among 2, 4, or 8 alternative stimuli. A single run was defined as 24 successive stimulus presentations; 48 different runs were available from the programer. The tests were carried out in groups of four runs. Each of the alternative stimuli occurred equally often in every run. The programer selected among eight durations, eight intensities, or eight combinations of frequency, intensity, and duration.

Each *S* was equipped with a response box which contained eight push buttons, eight corresponding lights, and a set of bulbs which marked the temporal sequence of Ready, Vote, Score. The responses of the 6 *S*s were summed directly into counters so that the complete stimulus-response matrix was available in real time. Only a single response button could be pressed at any one time. The correct response was indicated to *S*s at the end of each temporal sequence. In order to encourage individual performance, the number of correct responses per *S* was obtained and served as the basis of a monetary incentive system.

[2] Experiment V is not presented because an adequate amount of data is not available.

In each experiment *Ss* were given extensive practice with the stimulus objects before data were recorded. Each study was completed before the next was begun.

Treatment of results.—For each stimulus-response matrix, the following quantities were calculated: (*a*) the observed conditional response probability per entry; (*b*) when applicable, the derived conditional response probability per entry; (*c*) the mean difference (without regard to sign) between the derived and observed probabilities; (*d*) the percentage of occurrences in which the derived and observed probabilities differ by a proportion greater than .10; and (*e*) the mean difference (with regard to sign) between the derived and observed probabilities for the diagonal matrix entries (the correct responses).

Contextual Effects in Single-Dimensional Ensembles

Three experiments (III, IX, and VII) were specifically designed to examine contextual effects in single-dimensional ensembles. The absence of systematic contextual effects tends to support the CRR, whereas the presence of such effects tends to qualify the rule.

Experiment III: Ensemble Range

Experiment III was designed to study the derived confusability of a pair of stimulus objects as a function of the range of the stimulus ensemble.

Method.—A total of 21 stimulus ensembles, each consisting of four durations of a 1000-cps tone, were employed. The ensembles are schematically illustrated in Table 2. Each section of the table illustrates 7 ensembles in which two common stimulus objects are employed, and in which the range of the ensemble increases on successive rows. The 16 steps ranged from .097 sec. to 1.29 sec.

TABLE 2

Schematic Plan of Stimulus Ensembles Used in Exp. III (Duration) and IX (Intensity)

Ensemble	Stimulus Objects																Derived Confusability (%)	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Exp. III	Exp. IX
1	o	o	x	x													36	33
2	o	o			x	x											19	25
3	o	o					x	x									31	
4	o	o							x	x							30	28
5	o	o									x	x					22	
6	o	o											x	x			24	
7	o	o													x	x	35	32
				4				8				12				16		
8							x	o	o	x							25	36
9						x		o	o		x						24	
10					x			o	o			x					26	
11				x				o	o				x				23	
12			x					o	o					x			21	
13		x						o	o						x		24	
14	x							o	o							x	32	
				4				8				12				16		
15													x	x	o	o	21	36
16											x	x			o	o	21	
17									x	x					o	o	23	
18							x	x							o	o	24	30
19					x	x									o	o	21	
20			x	x											o	o	17	
21	x	x													o	o	20	27
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16		

Note.—In each section, two (open circles) of the four stimulus objects were present in all ensembles. The columns at the right give the derived confusability score, when available, for the common pair of stimulus objects.

TABLE 3

NUMBER OF OBSERVATIONS OF EACH STIMULUS IN EACH MATRIX AND NUMBER OF MATRICES (IN PARENTHESES) IN EACH EXPERIMENT

Experiment	Matrix Size		
	8 × 8	4 × 4	2 × 2
Contextual effects			
III (duration)		576(21)	576(3)
IX (intensity)		576(8)	576(3)
VII (intensity)[a]			
Series A	576	576(2)	576(5)
Series B	576	576(4)	576(10)
Series C	576	576(4)	576(8)
Series D	576	576(3)	576(10)
Single dimensional			
I (duration)	864	864(12)	864(14)
IV (intensity)	864	864(11)	864(14)
Multidimensional			
II (coarse)	864	864(13)	864(14)
VIII (fine)	576	576(8)	576(9)
VI (redundant)	432	864(10)	

[a] Also designed as a single-dimensional experiment.

in approximately equal logarithmic steps. The sound intensity was 80 db. spl. The observed confusability of each of the common pairs, i.e., 1, 2 in the top section; 8, 9 in the middle section; and 15, 16 in the lower section, was obtained in separate 2 × 2 tests.

Within each of the three sections, the individual ensembles of four stimuli were examined in a scrambled order. All observations associated with a given ensemble were completed before exposure of another ensemble. The number of observations of each stimulus in each 4 × 4 and each 2 × 2 confusion matrix is given in Table 3. A recognized limitation of the procedure is the possibility of inadequate counterbalancing of effects distributed in time (learning, motivation, morning vs. afternoon runs).

Results.—The derived confusability for the common stimulus pair of each ensemble is presented on the right side of Table 2. No clearly discernible trend within the large variability of the scores is observed as the range is increased.

The observed confusability (from the separate 2 × 2 tests) was 14, 25, and 24% for the common pair in the top, middle, and lower section, respectively, of Table 2.

Experiment IX: Ensemble Range

Method.—Experiment IX was identical with Exp. III except that: successive steps of intensity, spaced 1 db. apart, were employed; the duration was .32 sec.; Stimulus 1 of Table 2 corresponded to 70 db. spl; and fewer ensembles were examined.

Results.—The derived confusability scores are presented in the second right-hand column of Table 2. The observed confusability for the three common pairs was 39, 35, and 32%, respectively. As in Exp. III, the derived confusability scores do not systematically increase with an increase in ensemble range.

The RBM suggests that the derived confusability of a given stimulus pair should systematically increase as a function of the range of the stimulus ensemble. Experiments III and IX failed to observe this change, although the result must be qualified by the apparently large variability of the scores between successive conditions. Despite the large variability, the CRR describes the results reasonably well in that the derived and observed confusability are not systematically different.

Experiment VII: Ensemble Spacing

Experiment VII was designed to examine the joint effects of interstimulus spacing and ensemble range upon the derived confusability of particular pairs of stimulus objects. It also permitted an examination of the relative confusability of any pair of stimuli (and responses) as a function of the number of intermediate stimuli (and responses) between the pair.

Method.—Four interstimulus spacings were employed. They are illustrated in Table 4. Successive steps in the intensity of a 1000-cps tone, 0.32 sec. in duration, were separated by 4, 2, 1, and 0.5 db. for Series A, B, C, and D, respectively. A relative level of 0 db. represents a sound level of 70 db. spl. The

TABLE 4

SCHEMATIC ILLUSTRATION OF THE STIMULUS SPACINGS IN EXP. VII

Series	Relative Sound Level in db.									
	0		4		8	12	16	20	24	28
A (4 db.)	o		o		o	o	o	o	o	o
B (2 db.)	o	o	o	o	o	o				
C (1 db.)	o	o o o	o	o o o						
D (.5 db.)	o	o o o o o o o								

Note.—The open circles indicate the stimuli employed in each ensemble.

response label "1" was employed to designate the lowest intensity of a series; the response "8" to designate the highest intensity. Within a given series the labels employed in the 8 × 8, 4 × 4, and 2 × 2 tests were identical. However, from series to series, a given sound level was assigned a different label, e.g., a relative level of 4 db. was labeled "2" in Series A, "3" in Series B, and "5" in Series C.

Table 4 shows that the first two steps of Series A (hereafter written as A-1,2), Steps B-1,3; and Steps C-1,5 represent the same pair of sound intensities and that D-1,8 nearly represent the same pair of sound intensities. Similarly, A-1,3 and B-1,5 represent another pair of sound intensities. The number of intermediate stimuli between a given stimulus pair increases with a decrease in ensemble range. Thus, there are no intermediate stimuli (and their corresponding responses) between A-1,2; one between B-1,3; three between C-1,5; and six between D-1,8.

Each series was run in two separate blocks to permit a gross randomization of matrix size. In each block one-half of the data for the 8 × 8, 4 × 4, and 2 × 2 matrices of a given series was obtained in random order. The total number of responses to each stimulus in each matrix is given in Table 3.

Results.—The results are summarized in Table 5. The first column represents the distance in decibels between the two members of the stimulus pair under examination. The second column locates the stimuli in Table 4. The next four columns present, when available, the derived confusability, as determined from the 8 × 8 matrix (without parentheses), and the observed confusability, as determined from the 2 × 2 matrix (within parentheses). All entries in a given row represent the same stimulus pair, except those marked with an "a" in which the spacing has been shortened by 1 step-unit (i.e., 1.0 db. in C and 0.5 db. in D).

To be noted first is the lack of systematic differences as a function of the "label" assigned in the 2 × 2 tests (compare the entries within parentheses across a given line of Table 5). Stated simply, if *S* is faced with an ensemble of only two events at relative labels of 0 and 2, there is little difference if these levels represent Steps 1 and 5 in Series D, Steps 1 and 3 in Series C, or Steps 1 and 2 in Series B.

The main point of Table 5 is that the derived confusability scores decrease as the range of the initial 8 × 8

TABLE 5

OBSERVED (IN PARENTHESES) AND DERIVED CONFUSABILITY SCORES OF SELECTED STIMULUS PAIRS IN EXP. VII

Stimulus Spacing (db.)	Relative db.	Series			
		A	B	C	D
8	0–8	5	2	0[a]	
	4–12	7	1		
4	0–4	15 (6)	8 (4)	4	2[a]
	4–8	19 (7)	10 (7)	4[a]	
	8–12	18 (8)	10 (7)		
2	0–2		19 (15)	14 (15)	13 (18)
	2–4		17	13 (14)	14[a]
	4–6		20 (20)	12 (19)	
1	0–1			21 (27)	28 (26)
	1–2			21	23 (28)
	2–3			23 (24)	23 (24)

[a] Entries of ensembles shortened by one step interval.

ensemble decreases, i.e., the entries without parentheses decrease across a given line in Table 5. Further, the systematic change in derived confusability with the range of the ensemble decreases as the spacing between the stimulus pair decreases. The change in derived confusability from Series A toward D is greatest with the 8-db. spacing and is negligible with the 1-db. spacing.

And finally, the derived confusability tends to be systematically higher than the observed confusability for the widely spaced series (Series A and B) and tends to be systematically lower for the closely spaced series (Series C and D). The greatest systematic differences, about 10% on the average, are associated with the widely spaced series (Series A).

The ensembles of Table 4 confound the total range of the ensemble with the number of intermediate stimuli (and responses). Thus, while Ensemble C has more intermediate steps than Ensemble A associated with the same stimulus pair, Ensemble C subtends a smaller range than Ensemble A. A series of 4×4 tests substantiated the major findings of Table 5, but, again, ensemble range was confounded with the number of intermediate stimuli.

The systematic changes of Table 5 appear to be due to a confusion among alternative responses rather than among alternative stimuli. Response confusion seems to be particularly evident in the widely spaced series (A and B) when the proximity of the response labels is close. For example, in 8×8 tests involving the relative sound levels of 0 and 4 db., the response labels "1" and "2" (Series A) may be confused by *S*, but the labels "1" and "5" (Series C) are less likely to be confused. In the closely spaced series, presumably confusions among stimuli overwhelm the response effect.[3]

[3] An informational analysis of Exp. VII indicated that the amount of information achieved with each stimulus series (relative to the maximum possible transmitted information) is independent of both the spacing within each stimulus series and the number of stimulus objects. The amount of transmitted information seems to be determined only by the range subtended by the stimulus series.

The CRR and Single-Dimensional Ensembles

Three experiments (I, IV, VII) were designed to evaluate the CRR with stimulus objects varying along a single auditory dimension. The adequacy of the rule is tested primarily by comparing confusability scores obtained from 4×4 and 2×2 matrices with those derived from 8×8 matrices. The general method of the studies is the same as that previously reported.

Experiment I: Duration

Method.—The stimulus ensemble consisted of eight durations (.185-.651 sec.), equally spaced on a logarithmic scale, of a 1000-cps tone at a sound intensity of 80 db. Half of the 8×8 responses were obtained at the beginning, and half at the end, of the experiment. The 4×4 data were collected before the 2×2 data. The number of different 4×4 and 2×2 matrices examined and the number of observations associated with each matrix entry are given in Table 3.

Results.—Figure 1 (4×4) presents a comparison of the derived and observed response probabilities of the individual entries of the 4×4 matrices. The abscissa is the predicted proportions—the derived probabilities—from the 8×8 matrix; the ordinate is the obtained proportions—the observed probabilities—from the 4×4 tests. The CRR states that the predicted and obtained proportions should fall exactly on the 45° line of Fig. 1. Except for random variations, the rule makes excellent predictions.

A summary of the entries of Fig. 1 is presented in Table 6. The first

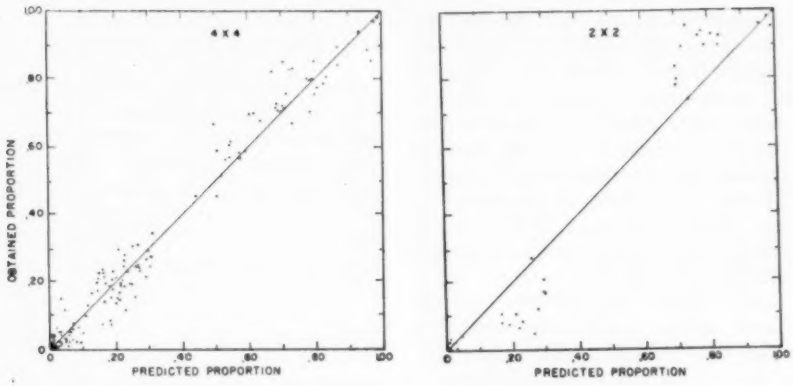

FIG. 1. Predicted and obtained response proportions of the individual 4×4 and 2×2 entries in Exp. I.

column gives the size of the matrix: 8×8, 4×4, or 2×2. The second column is the mean proportion of obtained correct responses. The third column is the mean of the differences (sign ignored) between the predicted and obtained proportions. The last column gives the mean systematic (sign not ignored) difference between the obtained and predicted proportions of the correct responses. A positive entry indicates that the obtained proportion was greater than the predicted proportion. In addition the obtained proportion correct scores of the 4×4 and 2×2 matrices were rank ordered, and then, on the basis of the ranks, the 4×4 and 2×2 matrices were divided equally into high and low proportion groups (first and second half, Table 6). The various difference measures were then calculated for the separate groups.

The entries of Table 6 (Exp. I) confirm the excellent CRR predictions of the 4×4 matrices shown in Fig. 1. The mean of absolute differences between the predicted and obtained scores is .028, and 4.2% (not given in Table 6) of the differences, on the average, are larger than .10, the arbitrary criterion used by Clarke (1957). The mean of the systematic differences is +.014, indicating that the obtained proportions are slightly larger than the predicted proportions, i.e., the derived confusability is greater than the observed confusability.[4] A comparison of the difference measures of the two halves of the 4×4 matrices indicates that somewhat better predictions are obtained with the matrices of the first half.

Table 6 and Fig. 1 also show that the CRR provides questionable predictions of the 2×2 matrices. The mean of the absolute differences is .082 and the average percentage of differences greater than .10 is 28.6. The mean systematic difference (+.052) suggests that the derived confusability is again greater than the observed confusability. This finding is graphically illustrated by

[4] In calculating the systematic differences for a given matrix, it was more convenient to use the proportion of correct responses than the proportion of incorrect responses. Confusability scores can be obtained by subtracting the proportion of correct responses from 1.0.

TABLE 6

SUMMARY OF RESULTS OF SINGLE-DIMENSIONAL AND MULTIDIMENSIONAL EXPERIMENTS

Matrix Size	Mean Prop. Correct	Mean |Diff.|	Mean Diff. of Diagonal Entries
Experiment I: Duration			
8 × 8	.568		
4 × 4	.746	.028	+.014
1st half	.784	.022	−.003
2nd half	.707	.034	+.032
2 × 2	.933	.082	+.052
1st half	.992	.057	.000
2nd half	.874	.107	+.105
Experiment IV: Intensity			
8 × 8	.713		
4 × 4	.831	.025	+.013
1st half	.860	.023	+.001
2nd half	.802	.027	+.025
2 × 2	.971	.043	+.042
1st half	.998	.003	.000
2nd half	.944	.083	+.084
Experiment II: Coarse Spacing			
8 × 8	.794		
4 × 4	.908	.020	+.020
2 × 2	.986	.023	+.018
Experiment VIII: Fine Spacing			
8 × 8	.485		
4 × 4	.666	.028	+.002
2 × 2	.846	.038	−.004
Experiment VI: Redundant Objects			
8 × 8	.834		
4 × 4	.947	.018	+.028

the clustering of the entries on one side of the diagonal in Fig. 1. The difference measures are especially large in the matrices with the lower proportion of correct responses (Table 6). The accuracy of prediction thus seems to be strongly dependent upon the obtained proportion of correct responses. The largest differences are always found in matrices with adjacent stimulus (and response) objects, e.g., 1,2; 5,6.

Experiment IV: Intensity

Method.—Experiment IV was identical with Exp. I except that an ensemble of eight intensities, spaced 3 db. apart, was examined. The lowest sound level was 70 db. spl. The frequency of the tone was 1000 cps and the duration was .32 sec. Table 3 gives the number of observations of each stimulus in the 8 × 8, 4 × 4, and 2 × 2 matrices.

Results.—Figure 2 shows a close correspondence between the obtained and predicted response probabilities for the 4 × 4 matrices. The mean of the absolute differences is .025 (Table 6) and the mean percentage of differences greater than .10 is 2.8. An analysis of the systematic differences between the predicted and obtained probabilities for the correct responses produced a mean difference of +.013. Thus, all the difference measures support the accuracy of the CRR predictions shown in Fig. 2. A comparison of the difference scores for the first and second halves of the 4 × 4 matrices indicates that the CRR predicts equally well the entries of matrices with high and lower proportion correct scores. This confirms a similar result obtained in Exp. I (Duration).

The predictions of the CRR are much poorer for the 2 × 2 data (Fig. 2 and Table 6). The mean absolute difference is .043; 21.5% of the differences exceed .10, on the average; and the mean of the systematic differences for the correct responses is +.042. An analysis of the two halves of the matrices (Table 6) indicates that the difference measures for the matrices with high proportion correct scores approach zero while those of the second half are larger. As in Exp. I, the largest differences are associated with matrices made up

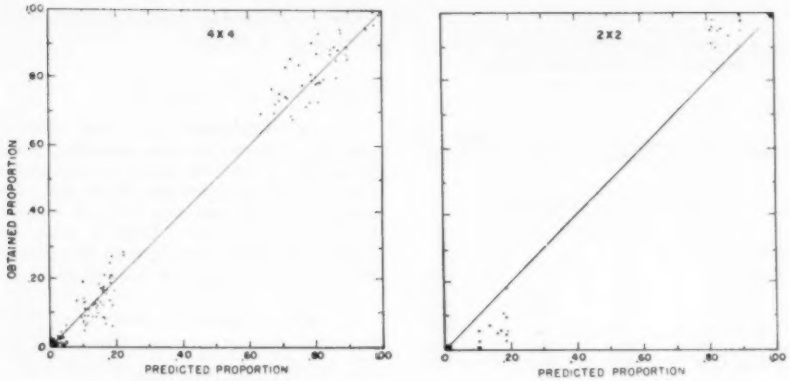

FIG. 2. Predicted and obtained response proportions of the individual 4×4 and 2×2 entries in Exp. IV.

of adjacent stimulus (and response) objects.

Experiment VII: Intensity Spacing

Method.—The method of Exp. VII was presented in the section concerned with contextual effects. Only the results relevant to the CRR will be presented here.

Results.—Table 7 summarizes the results for the four series of Exp. VII. Its organization is identical with that of Table 6. The mean difference measures for the 4×4 matrices demonstrate reasonably good CRR prediction. However, the mean of the diagonal differences (systematic effects) indicates that the rule underestimates the mean predicted proportion of correct responses for the widely spaced series (A and B) and overestimates those for the closely spaced series (C and D).

Inspection of the results of the 2×2 tests (Table 7) suggests that the CRR predictions are again reasonably good, with the possible exception of Series A (the widest ensemble.) The difference measures of Series A are all larger than those of the other series. Again, it should be noted that the largest differences are observed with matrices containing adjacent stimulus (and response) objects.

Discussion

The serious failures of the CRR are associated with adjacent steps where

TABLE 7

SUMMARY OF RESULTS OF EXP. VII (INTENSITY SPACING)

Series	Mean Prop. Correct	Mean |Diff.|	Mean Diff. of Diagonal Entries
	8×8		
A (4 db.)	.746		
B (2 db.)	.597		
C (1 db.)	.527		
D (0.5 db.)	.375		
	4×4		
A	.852	.034	+.053
B	.748	.025	+.006
C	.629	.041	−.068
D	.529	.042	−.006
	2×2		
A	.939	.088	+.084
B	.904	.034	+.022
C	.803	.051	−.030
D	.756	.054	−.029

the stimulus (and response) spacing is wide. In these instances the predicted proportions are seriously underestimated relative to the obtained proportions. The result is not unreasonable when examined in terms of stimulus and response confusions. A pair of stimulus objects, separated by a large stimulus difference, may be easily discriminated in a 2 × 2 test. However, within an 8 × 8 matrix, the same stimulus objects and their associated responses may be thoroughly confused. Presumably, the confusion is primarily between responses rather than between stimuli. The confusion among responses is most clearly evident in Series A of Exp. VII. Since the stimuli of Series A are more widely spaced than those of the other series, it seems inappropriate to assign the relatively large differences of Series A (especially those of the 2 × 2 data) to stimulus confusion. The other series of Exp. VII do not show such effects presumably because confusions among the stimuli of these series overwhelm any response effect. Furthermore, the overestimation of the predicted proportions in the closely spaced series (C and D) of Exp. VII suggests that the effect of stimulus confusions not only overwhelms but tends to oppose the response effect.

The importance of interstimulus spacing is shown more clearly in Fig. 3 which provides a further analysis of Exp. I.

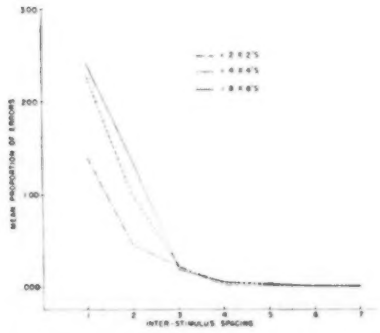

FIG. 3. Confusability scores in Exp. I as a function of interstimulus distance.

Derived confusability scores from the 8 × 8 matrix and from appropriate 4 × 4 matrices are plotted along with observed confusability scores from 2 × 2 matrices. The abscissa is the spacing between the members of selected stimulus pairs (and their responses) within a matrix; the ordinate is the mean proportion of incorrect responses (confusability score) for all pairs with the same interstimulus (and response) spacing, e.g., the Pairs 1 and 4 and 2 and 5 have the same spacing. The main feature of Fig. 3 is that the derived scores (from 8 × 8 and 4 × 4 matrices) are systematically higher than the observed scores (from 2 × 2 matrices) only when the interstimulus (and response) spacing is small, i.e., when the stimulus (and response) objects are adjacent or nearly adjacent. A similar analysis of Exp. IV also showed that the largest differences between observed and derived confusability occur in matrices with adjacent items. That the differences are due to confusion among responses rather than among stimuli is suggested by the widely spaced ensembles used in the experiments.

Finally, the accuracy of the CRR predictions must necessarily be a function of the relative performance level of *Ss*. Excellent predictions are always obtained when the proportion of correct responses approaches 1.00. A comparison of the results given in the first half and the second half of Table 6 (Exp. I and IV) shows that in every case the smallest differences are associated with matrices in which a relatively high proportion of correct responses was obtained.

THE CRR AND MULTIDIMENSIONAL ENSEMBLES

The final experiments (II, VIII, VI) were designed to test the CRR with stimulus objects which vary along three stimulus dimensions. These ensembles represent test conditions between the single-dimensional displays described earlier in the present report and complex displays such as

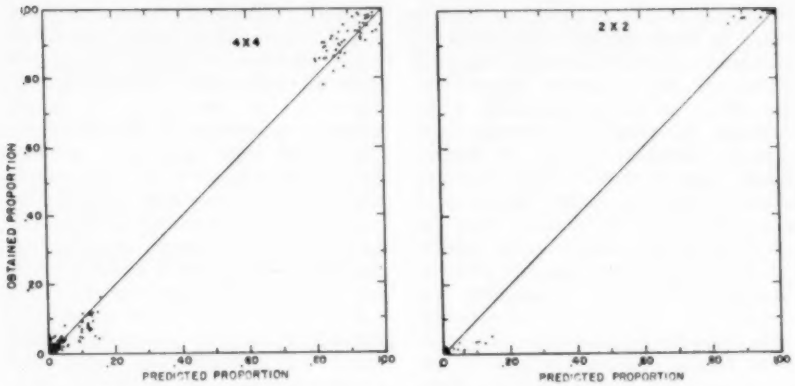

FIG. 4. Predicted and obtained response proportions of the individual 4 × 4 and 2 × 2 entries in Exp. II.

speech. The experiments examine CRR predictions as a function of interstimulus spacing and stimulus redundancy. The general method of the studies is the same as that in the single-dimensional experiments.

Experiment II: Coarse Multidimensional

Method.—An ensemble of eight stimulus objects was constructed from the combination of two tonal frequencies (1000 or 1025 cps), two durations (.255 or .367 sec.), and two sound levels (80 or 84 db.). For example, Stimulus 1 consisted of a tone 1000 cps in frequency, .255 sec. in duration, and 80 db. in intensity. Since the two values of each dimension were rather widely spaced, the ensemble is described as coarse. In all of the present multidimensional experiments, *Ss*' push buttons were labeled with respect to all three dimensions. For example, the button for Stimulus 1 was labeled: low, short, soft. Table 3 summarizes the number of responses obtained in the various matrices.

Results.—Figure 4 presents a comparison of the predicted and obtained proportions of the 4 × 4 matrices.

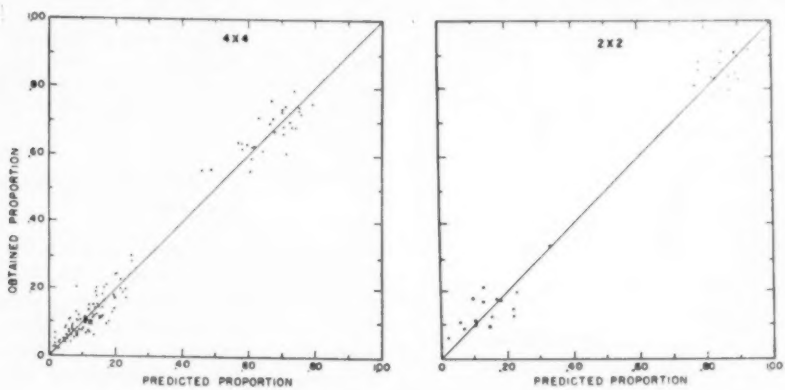

FIG. 5. Predicted and obtained response proportions of the individual 4 × 4 and 2 × 2 entries in Exp. VIII.

Although some of the data show systematic deviation from the 45° line, the results are generally well described by the CRR. The mean difference measures (Table 6) are all small. In both the 4 × 4 and 2 × 2 matrices of the multidimensional studies, the mean percentage of differences exceeding .10 was less than two. The results of the 2 × 2 matrices (Fig. 4 and Table 6) are also fairly well described by the CRR.

Experiment VIII: Fine Multidimensional

Method.—The eight stimulus objects varied with respect to: a frequency of 1000 or 1006 cps, a duration of .320 or .367 sec., and an intensity of 80 or 81 db. Experiments II and VIII were otherwise identical except that data collection followed the block system (method of counterbalancing) of Exp. VII. The number of presentations of each stimulus in each ensemble is given in Table 3.

Results.—The close correspondence between the obtained and predicted proportions of the 4 × 4 matrices (Fig. 5) is quantitatively confirmed by the low values of the difference measures presented in Table 6. Comparable high agreement between the obtained and predicted proportions was also the case for the data of the 2 × 2 matrices (Fig. 5 and Table 6).

Experiment VI: Redundant Multidimensional

Method.—In the final multidimensional study, the stimulus ensemble was constructed from the combination of two frequencies (1000 and 1025 cps), four durations (.237, .320, .450, and .651 sec.), and four intensities (86, 89, 92, and 95 db.). Intensity and duration were made redundant by the use of four values for each dimension. But there was not a one to one relation between a particular sound level and a particular duration, e.g., the 86-db. level was paired with both the .237- and .450-sec. durations. Data were collected only for the 8 × 8 matrix and 10 4 × 4 matrices (Table 3).

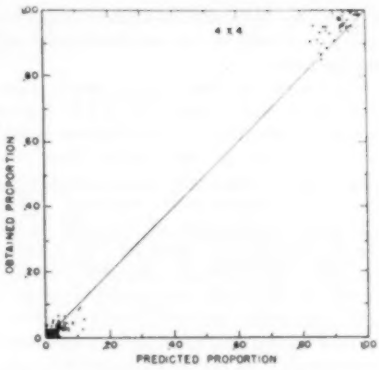

FIG. 6. Predicted and obtained response proportions of the individual 4 × 4 entries in Exp. VI.

Results.—The results are shown in Fig. 6 and Table 6. Although there is good agreement between the obtained and predicted proportions, there is a definite tendency for the points to deviate systematically from the 45° line. The extent of the deviations is more precisely indicated by the relatively large mean of diagonal differences (+.028).

Discussion

In all of the multidimensional experiments the CRR makes good or excellent predictions of the response probabilities of the various 4 × 4 and 2 × 2 matrices. Furthermore, the rule seems to hold under a variety of conditions in the multidimensional situation, e.g., coarseness and fineness of the display, the presence or absence of stimulus redundancy.

Perhaps a more important finding, however, is that the CRR predicts the response probabilities of the multidimensional matrices somewhat better than those of the single-dimensional matrices. For example, in eight of the single dimensional and in none of the multidimensional tests, the mean difference scores exceed .04. A similar analysis of the mean difference of diago-

nal entries (systematic effects) supports the superiority of the multidimensional predictions.

Although evidence has been marshalled to show that the CRR makes better predictions in multidimensional situations, it is also clear that the rule makes relatively good predictions for single-dimensional displays. The successful application of the rule in the single-dimensional cases of the present experiments is directly opposite Clarke's (1959) finding that responses to intensity and frequency stimuli cannot be described by the CRR. Clarke's difficulty with the rule, however, was due to strong contextual effects present in his data.[5]

Summary

Eight experiments were designed to evaluate a model—the constant-ratio rule—of sensory organization. The model was tested by comparing the obtained probabilities of 4×4 and 2×2 confusion matrices with the probabilities predicted from 8×8 matrices. The auditory dimensions of frequency, intensity, and duration were used singly and in combination to form single and multidimensional stimulus ensembles consisting of two, four, or eight tones. Six *S*s attempted to identify, one at a time, the stimulus objects of the ensembles.

The results indicate: (*a*) In general, the CRR makes excellent predictions of the response probabilities of the single and multidimensional matrices. The notable exceptions occur when predictions for 2×2 matrices are made from adjacent objects in widely spaced ensembles. (*b*) The CRR predicts the probabilities of the multidimensional matrices somewhat better than those of the single-dimensional ensembles.

[5] F. R. Clarke, personal communication, 1960.

References

Anderson, C. D. The constant-ratio rule as a predictor of confusions among visual stimuli of brief exposure duration. *USAF Cambridge Res. Cent. tech. Note*, 1959, No. TN 58–60. (ASTIA No. AD-160706).

Clarke, F. R. Constant-ratio rule for confusion matrices in speech communication. *J. Acoust. Soc. Amer.*, 1957, **29**, 715–720.

Clarke, F. R. Proportion of correct responses as a function of the number of stimulus-response alternatives. *J. Acoust. Soc. Amer.*, 1959, **31**, 835. (Abstract)

Clarke, F. R., & Anderson, C. D. Further test of the constant-ratio rule in speech communication. *J. Acoust. Soc. Amer.*, 1957, **12**, 1318–1320.

McGill, W. J. The slope of the loudness function: A puzzle. In H. Gulliksen & S. Messick (Eds.), *Psychological scaling*. New York: Wiley, 1960. Pp. 67–81.

Pollack, I., & Decker, L. Consonant confusions and the constant ratio rule. *Lang. Speech*, 1960, **3**, 1–6.

Torgerson, W. S. *Theory and methods of scaling*. New York: Wiley, 1958.

(Received November 12, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 2, 143–147

REACTION TIME TO STIMULI MASKED BY METACONTRAST [1]

ELIZABETH FEHRER AND DAVID RAAB

Brooklyn College

The experiment to be described below is concerned with the relation between (*a*) phenomenal characteristics and (*b*) an objective behavioral measure of metacontrast suppression. If a brief light flash is followed, after a suitable interval, by stimulation of adjacent retinal areas, its apparent brightness is reduced. The extent of such retroactive masking has been shown to be a U shaped function of stimulus onset asynchrony, with the maximum suppression occurring at asynchronies of 70 to 100 msec. (e.g., Alpern, 1953). At such asynchronies, the first stimulus may be phenomenally absent (Kolers & Rosner, 1960; Toch, 1956; Werner, 1935). It is not, however, without effect on perception since under these conditions there is an awareness of movement involving the second stimulus. (Such asynchronous stimulation of adjacent retinal areas is, of course, the condition for producing apparent movement.)

The problem of the present experiment was to determine whether the phenomenal masking of the first stimulus is associated with a decrease in its capacity to elicit a simple overt response. Specifically, we hoped to determine whether reaction time (RT) is correlated with the apparent brightness of a flash of light subjected to various degrees of metacontrast masking. The question becomes of especial interest at those stimulus onset asynchronies at which the first flash is phenomenally absent and only the masking stimuli (flashed second) are perceived.

The present problem is related to that raised in a previous study (Raab, Fehrer, & Hershenson, 1961) which dealt with the relation between RT and the apparent brightness of light flashes as contingent on their duration. It was shown that RT depended on stimulus characteristics rather than on the phenomenal appearance of the stimulus.

RT is a convenient measure of retroactive masking since the masked stimulus is the one presented first. If the RT chronoscope is started when the first stimulus is presented, it is possible to compare the RT to this stimulus when it is presented alone with the RT to it when it is followed by masking stimulation. Any lengthening of RT can then be attributed to the masking procedure.

METHOD

Apparatus

The stimulus lights and *S* were located in one room; *E*'s control panel and the RT chronoscope were in another room.

The stimulus lamps were placed behind a 6 × 2 in. rectangular box divided by two partitions into three 2 × 2 in. square cells arranged horizontally (see Fig. 1). The top, bottom, sides, and septa of the box were made of sheet aluminum, $\frac{1}{32}$ in. thick. The front of the array (as seen by *S*) was covered with a sheet of diffusing tissue pressed against the cell walls by a 18.5 × 8.5 in. pane of $\frac{1}{4}$-in. polished plate glass mounted in a 2-in. wooden frame. The entire assembly was placed at eye level and was viewed binocularly by *S* from a distance of 15 ft. The horizontal visual angle of the array was 2°. For foveal stimulation, two pin-point red lights were placed at the locations marked F (Fig. 1),

[1] This research was supported by Grant G-6456 from the National Science Foundation.

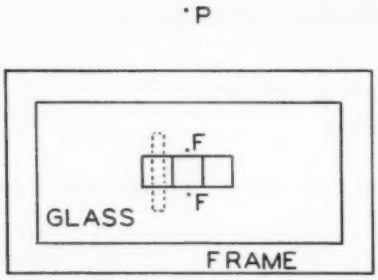

FIG. 1. The stimulus array as seen by *S*. (The three light cells are covered by a sheet of diffusing tissue. The position of the fluorescent tube in each cell is indicated by the dashed outline in the left-hand cell. F and P are fixation points. For dimensions, see text.)

and *S* was instructed to fixate between them. Autokinetic movement of these fixation lights in the otherwise dark test room was eliminated by dimly illuminating the frame surrounding the plate glass. This illumination was too faint to be measured. The frame did not become visible to *S* until after 5 min. of dark adaptation.

For peripheral stimulation, a single fixation light, P (Fig. 1), was placed 8.75 in. (i.e., 2° 50′) above the top center of the central light cell.

A cold cathode fluorescent lamp was placed at the rear of each of the three light cells with phosphored portion 7¼ in. behind the diffusing tissue viewed by *S*. The lamps are of the mercury vapor type and are coated with a calcium halo phosphor. Luminance, measured at the diffusing tissue, was approximately 18 ft-L.

The lamps were pulsed by a "hard tube" high-voltage circuit. The vacuum tubes in series with each lamp were driven by a Tektronix 161 pulse generator. For all the work reported here, the lamp in the center cell was pulsed by one circuit and the two flanking cells were flashed together by an independent circuit. Regulated power supplies were employed throughout. When flashed, the lamps provide reasonably rectangular pulses of light. The chief departure occurs at offset and results from phosphor persistence.

Only two durations of flash, 5 and 50 msec., were used in this experiment. Since the light output from a lamp increases slightly with durations beyond 5 msec., the luminances of the two pulses at offset were not equal. The 50-msec. pulse reached a level 0.076 log units above that of the 5-msec. flash.

Timing of the stimulus pulses was accomplished by means of a set of Tektronix pulse and wave form generators. All critical timing voltages were regulated and the times themselves were monitored on a Westport WE-210 counter.

RT trials were presented once every 15 sec. Each trial began with a 1-sec. warning tone. The stimulus pattern was presented (and the counter started) 2.9, 3.2, or 3.5 sec. after the onset of the tone. The three foreperiods were switch selected and were presented in random order. The counter was stopped when *S* depressed a normally closed telegraph key.

RTs to the center light alone and to the center-plus-flanks could be read directly from the counter. RTs to the flanks alone were computed by subtracting the stimulus asynchrony from the counter readings.

Procedure

Preliminary exploration indicated that good metacontrast suppression could be obtained with various durations of the center and flanking lights. It was arbitrarily decided to limit the present investigation to the following combinations: (*a*) center, 50 msec.; flanks, 50 msec. (50-50-50); (*b*) center, 5 msec.; flanks, 50 msec. (50-5-50). Each combination was studied in foveal and in peripheral vision.

Preliminary.—Systematic variation of stimulus onset asynchrony (Δt) yielded the following perceptual correlates of asynchrony for the 50-50-50 sequence with foveal vision. The first perceptible darkening of the center light (as compared with the flanks) appears at a Δt of 4 to 6 msec. The darkening progresses with increases in Δt to about 35 msec. At this asynchrony, the center field is no longer homogeneous in color nor phenomenally a square, but appears as a faint, shapeless flash of light in the center of an extremely dark field between the flanks. Maximum suppression occurs with a Δt of approximately 75 msec. With increases in Δt beyond about 90 to 100 msec., depending on *S*, the center flash brightens again. At approximately 120 msec., it appears as a square preceding the flanks and of the same brightness as the flanks. Its apparent duration, however, is less than that of the flanks. With a Δt of about 150 msec., both brightness and duration of the center seem the same as those of the flanks.

With the 50-5-50 lights in foveal vision,

the center square does not appear homogeneous even with a Δt of zero. Instead, only its center appears light. Maximum darkening and reappearance occur at about the same Δt's as for the 50-50-50 lights.

In peripheral vision, the phenomenal correlates of asynchrony differed from those described above for foveal vision only in that the apparent darkening of the center light was more pronounced than with foveal vision. With Δt's from 50 to 100 msec., the entire center square area appeared completely dark.

On the basis of the above findings, the following asynchronies were selected for systematic study: 0, 10, 25, 50, and 75 msec. It was assumed that, if RT parallels apparent brightness of the center light, RT, measured from the onset of the center light, should increase with Δt, reaching a maximum at a Δt of 75 msec., and that here, especially in peripheral vision, it might well be the same as RT to the flanks.

Foveal vision.—The authors and a graduate student served as *S*s. All had had extensive previous experience with RT. Each *S* served in 50 sessions, 25 with the 50-50-50 durations and 25 with the 50-5-50. Of the 25 sessions at a given duration of the center light, 5 were assigned to each of the 5 stimulus onset asynchronies.

Thirty-six RTs were recorded at each session, 12 to the center light alone, 12 to the flanks alone, and 12 to the combination. Random presentation prevented *S* from knowing which stimulus pattern would be presented at a given trial. Foreperiod duration was also randomized.

The longest and the shortest RT of each 12 were eliminated before the three means for each session were computed. Since there were five sessions for each asynchrony, the means for each *S* are based on 50 RTs.

Two sessions were run on each experimental day, one with the 50-50-50 and one with the 50-5-50 durations. Sessions at the various Δt's followed each other in random order.

After 5 min. of dark adaptation, three practice trials were given. Then the recorded trials were run. A 1-min. rest period separated each 18 trials.

Peripheral vision.—For the 50-50-50 durations, two graduate students and one undergraduate major served as *S*s. For the 50-5-50, one author and one undergraduate major were *S*s. Each served in 20 sessions, 5 each at the following four Δt's: 0, 10, 25, and 75 msec. The procedure was otherwise identical with that employed for foveal vision.

TABLE 1

MEAN REACTION TIMES WITH FOVEAL STIMULATION

Δt in Msec.	50-50-50			50-5-50		
	Center	Flanks	All	Center	Flanks	All
0	165.9	163.3	161.1	167.3	162.3	162.6
10	164.6	162.3	161.9	166.2	162.6	161.3
25	167.1	163.4	164.9	169.6	166.6	167.1
50	167.2	166.6	167.5	166.5	162.5	166.5
75	162.4	160.3	164.5	165.4	161.9	165.1
Mean	165.4	163.2		167.0	163.2	

RESULTS AND DISCUSSION

Since all *S*s showed the same trends, only the group means are presented in Tables 1 and 2. The individual *SD*s were of the order of 12 msec. and did not differ with stimulus conditions. The data in both tables show that RT to the center light was never *increased* by the addition of the masking stimuli. A series of *t* tests of the difference between each individual center and center-plus-flanks means showed that in no case was the RT to center-plus-flanks reliably longer. It follows, therefore, that decreases in apparent brightness produced by the metacontrast procedure are not reflected in RT increases. This is true even for peripheral vision where, with a Δt of 75 msec., the center flash is phenomenally absent. Under all our conditions, RT seemed to be determined only by the *stimulus*

TABLE 2

MEAN REACTION TIMES WITH PERIPHERAL STIMULATION

Δt in Msec.	50-50-50			50-5-50		
	Center	Flanks	All	Center	Flanks	All
0	159.6	155.2	153.4	167.6	160.2	160.9
10	162.3	155.6	158.9	164.4	160.0	163.5
25	163.2	160.5	164.1	164.8	159.6	164.3
75	161.8	158.1	161.5	165.7	160.0	165.0
Mean	161.8	157.3		165.6	159.9	

characteristics of the light flashed first.

The fact that the stimulus does evoke a normal RT even when completely "masked" is not, perhaps, entirely unexpected. Since the metacontrast procedure is also the procedure for producing phi movement, it is obvious that the light flashed first does affect the total perception. Even though the center light appears "blanked," its prior occurrence is a necessary condition for the perception of movement involving the flanks, and such movement was apparent to all our *Ss*. Our data have shown clearly that RT to a light is the same when (*a*) it is presented alone and perceived as a bright flash, and when (*b*) it is followed by later stimulation of adjacent retinal areas and is "perceived" only through the apparent movement of these lights.

Our results do not depend on knowledge of stimulus conditions. RTs of 6 naive *Ss* were measured to the center light alone and to the center followed after 75 msec. by the flanks. The 50-50-50 lights and peripheral fixation were used. Although the RTs were longer than those appearing in our tables (*Ss* were unpracticed), they were the same under both stimulus conditions. When these *Ss* were asked to describe what they saw, all reported seeing either a single square of light or two squares separated by a dark area about the size of one of the squares. Movement of the two squares was never spontaneously reported until at least 20 RT trials had been run.

The *Ss* were then tested to see whether they could learn to distinguish phenomenally between the appearance of the flanks alone and the center-plus-flanks. All but 1 learned to do so but not until 10 to 20 trials (with correction) had been run.

In other words, these *Ss*' RTs were from the beginning a reliable index of the presence of the "masked" center light whereas their phenomenal reports became adequate only after some specific training.

The data provide some information concerning the effect of stimulus duration on RT. The center stimulus was presented alone on one-third of the trials in each session. Since both 5-and 50-msec. flashes were employed, RTs to these can be compared directly. With foveal vision (Table 1), the mean RT to the 50-msec. flash was 165.4 msec.; the mean RT of the same *Ss* to the 5-msec. pulse was 167.0 msec. The absence of any marked difference in RT is in line with the results of a previous study (Raab et al., 1961) in which RT was found to be independent of stimulus durations between 10 and 500 msec. A similar comparison cannot be made with the present peripheral data since the 5- and 50-msec. flashes were presented to different *Ss*.

That stimulus size affects RT (e.g., Froeberg, 1907) is shown by our data. Reactions to the single 50-msec. center flash were slower than to the two 50-msec. flanks. Of the 15 possible individual comparisons between center and flank means for foveal vision (3 *Ss* $\times$ 5 asynchronies), 13 show faster RT to the flanks than to the center ($P = .01$). Of the 12 comparisons (3 *Ss* $\times$ 4 asynchronies) in peripheral vision, 11 show faster RT to the flanks ($P < .01$). For foveal vision, the overall mean of 165.4 msec. for the center light differs significantly ($P < .01$) from the mean of 163.2 for the flanks (Table 1, last row). The comparable difference for the peripheral data is also significant.

When the center and flanks are pulsed simultaneously ($\Delta t = 0$), the stimulus area is, in effect, three times that of the center alone. RT to this largest stimulus was in all cases faster than RT to the flanks.

It is interesting to note that RT to the center-plus-flanks is also faster than RT to the center alone when the

flanks are delayed by 10 msec. Of the 11 differences between individual means at this asynchrony (6 for foveal and 5 for peripheral vision), 10 show this effect ($P = .01$). There is some indication that this effect persists up to the 25-msec. Δt for foveal vision, since in the six possible comparisons under this condition, RT to the combined stimulation was faster in all cases than RT to the center alone. The finding of "facilitation" at these two asynchronies is the opposite of what would have been found if phenomenal brightness governs RT since, at these asynchronies, the center flash is appreciably darkened. The finding is similar to an intersensory facilitation of RT recently reported by Hershenson (1960).

Summary

This study was designed to determine whether metacontrast suppression of a light flash affects RT to the flash. Masking of a square target was achieved by subsequent flashing of two adjacent squares. Stimulus onset asynchronies were studied over a range from 0 to 75 msec. Phenomenal suppression of the first flash varied from none (0 asynchrony) to maximum (75 msec. asynchrony). With both foveal and peripheral vision, the results showed no effect of phenomenal brightness on RT. RT to the target, flashed alone, was never faster than RT to the target followed by masks. It was therefore the physical dimensions of the stimulus rather than its phenomenal characteristics that determined RT.

References

Alpern, M. Metacontrast. *J. Opt. Soc. Amer.*, 1953, **43**, 648–657.

Froeberg, S. The relation between the magnitude of stimulus and the time of reaction. *Arch. Psychol., NY*, 1907, **16**, 1–38.

Hershenson, M. Reaction time as a measure of intersensory facilitation. Unpublished master's thesis, Brooklyn College, 1960.

Kolers, P. A., & Rosner, B. S. On visual masking (metacontrast): Dichoptic observation. *Amer. J. Psychol.*, 1960, **73**, 2–21.

Raab, D., Fehrer, E., & Hershenson, M. Visual reaction time and the Broca-Sulzer phenomenon. *J. exp. Psychol.*, 1961, **61**, 193–199.

Toch, H. The perceptual elaboration of stroboscopic presentations. *Amer. J. Psychol.*, 1956, **69**, 345–358.

Werner, H. Studies on contour: I. Qualitative analysis. *Amer. J. Psychol.*, 1935, **47**, 40–64.

(Received November 15, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 2, 148–154

THE EFFECT OF POSTURAL FACTORS ON THE DISTRIBUTION OF TACTUAL SENSITIVITY AND THE ORGANIZATION OF TACTUAL-KINAESTHETIC SPACE[1]

JOSEPH H. McFARLAND, HEINZ WERNER, AND SEYMOUR WAPNER

Clark University

The distribution of tactual sensitivity has been widely held to correspond with the spatial distribution of touch receptors on the body: where there is a high density of receptors there is high sensitivity. There are three lines of evidence however, which question the plausibility of this strictly peripheral explanation of the distribution of tactual sensitivity first proposed by Weber (1846). First, on anatomical grounds it is questionable as to whether there is a strict one-to-one correspondence between the spatial distribution of touch receptors and the distribution of tactual sensitivity (Jenkins, 1951). Second, it is questionable as to whether high density of receptors is the sole condition for high sensitivity, since sensitivity depends on judgmental factors (Boring, 1942, pp. 480–483; Gibson, 1953). Third, it is questionable as to whether tactual sensitivity can be explained solely in terms of the tactual afferent system, since studies with cerebellar patients suggest that sensitivity also depends on the postural status of the organism (Halpern & Kugelmas, 1957).

The first aim of this paper is to consider further the question of whether tactual sensitivity can be explained solely in terms of the tactual afferent system by systematically investigating the effect of postural changes on sensitivity in normal adults.

The second aim of this paper is to generalize from previous findings in the visual modality to the tactual modality. Previously McFarland (1957) and McFarland, Werner, and Wapner (1960) have shown that changes in the pattern of proprioceptive stimulation from the oculomotor system lead to a redistribution of visual sensitivity. With increased convergence of the right eye during binocular viewing, there is an increase in acuity to the right of the focal point; conversely, with increased convergence of the left eye during binocular viewing, there is an increase in acuity to the left of the focal point. Concomitant with these changes in acuity there is a reorganization of visual space: with increased convergence of the right eye during binocular viewing, there is a shift of the apparent median plane (the straight-ahead) to the right of the objective median plane; conversely, with increased convergence of the left eye during binocular viewing, there is a shift of the apparent median plane to the left of the objective median plane (Bruell, 1953). The question to be investigated is whether there is a concomitant redistribution in tactual sensitivity and reorganization

[1] This investigation was supported by a USPHS grant, M-348, from the National Institute of Mental Health, United States Public Health Service. The authors wish to thank David Calverley who served as Research Assistant for some of the experiments.

in tactual-kinaesthetic space depending on postural changes.

Two experiments are reported: the first deals with the effect of postural changes on the distribution of tactual sensitivity and the second deals with the effect of postural changes on the organization of tactual-kinaesthetic space.

Experiment I

Generalizing from the earlier findings concerning a redistribution in visual sensitivity, the following hypothesis was formulated: tilt left increases sensitivity of the right palm compared to the left palm; tilt right increases sensitivity of the left palm compared to the right palm. Three sets of postural conditions—head tilt, supported body tilt, and unsupported body tilt—were used, rather than only one set, in order to increase the breadth of the findings with respect to the predicted relationship between postural changes and tactual sensitivity changes.

Method.—For each *S*, a tactual two-point threshold was obtained for left and right palms. Independent groups of *S*s were used for each set of postural conditions. Each set of postural conditions consisted of three positions: tilt 20° left (L), erect (E), and tilt 20° right (R). These variations within a set were arranged in a 3 × 3 replicated latin square. For each sequence, half the *S*s were tested first for left palm and then for right palm; the remaining half were tested first for right palm and then for left palm. Thus for each *S* a tactual two-point threshold was obtained for left and right palms under a set of postural conditions.

The important features of two aesthesiometers (Fig. 1) constructed for this experiment are as follows: (*a*) Two plastic needles (A) are mounted in sleeves set at right angles to a base (B). Set screws in the sleeves (C) permit independent variation of the height of the needles. Thus, the needles can be adjusted to compensate for irregularities of the palm's surface so that the needles strike the palm at the same time and with equal pressure. Total weight of the needles and base is 22 gm.

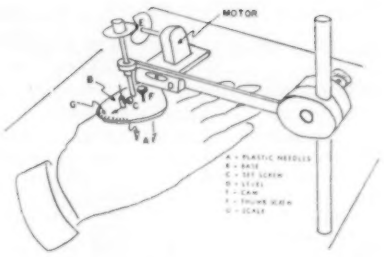

Fig. 1. Schema of aesthesiometer used to determine two-point threshold.

(*b*) The needles may be oriented by means of a level (D) so that they fall perpendicularly. (*c*) By means of a motor driven cam (E), the needles drop onto the palm, remain there for 2 sec., then are raised and after 23 sec. drop again. (*d*) Distance between the needles can be varied by means of a thumb screw and gear arrangement in the base (F). Distance between the needles in millimeter units can be read directly from a scale on the base (G).

Each S extended his forearms at right angles to his trunk resting the dorsal side of the hands on individually fashioned Plasticine moulds. The *S*'s arms and hands were kept in this position throughout the experimental session. With *S*'s eyes closed, a two-point threshold was determined by the method of limits using one ascending and one descending trial. In the ascending series the points were started 2 mm. apart. In the descending series the points were started 20 mm. apart. Successive trials in each series differed by 2-mm. steps. The *S* was restricted to two categories of judgment: "one-point" or "two-points"; he was instructed to use the category "two-points" only when he clearly experienced two points. There was a 23-sec. interval between each trial in a series. After a threshold was determined for one palm, this procedure was repeated *immediately* for the other palm.

In the analysis, the log of the difference between thresholds for left minus right palms plus a constant was utilized.

In Cond. 1 (Head Tilt) *S*, with body erect, assumed three head positions: L, E, and R. A scale behind *S* assured that he maintained the required head position. Hands and arms were kept in the same position relative to *S*'s body for all positions. Twenty-four *S*s were tested. In Cond. 2 (Supported Body Tilt) *S* stood in an open box (padded with foam rubber to keep him straight) which could be

TABLE 1

EXPERIMENT I: ANALYSES OF VARIANCE FOR LOG THRESHOLD DIFFERENCE SCORES BETWEEN LEFT AND RIGHT PALMS

Source	1. Head Tilt			2. Supported Body Tilt			3. Unsupported Body Tilt		
	df	*MS*	*F*	*df*	*MS*	*F*	*df*	*MS*	*F*
Ss	23	.0743	<1.00	11	.0331	1.08	23	.0104	<1.00
Sequence (S)	2	.0560	<1.00	2	.0440	1.43	2	.0030	<1.00
Ss within S	21	.0761	—	9	.0307	—	21	.0111	—
Order	2	.0065	<1.00	2	.0105	1.35	2	.0020	<1.00
Tilt	2	.0820	3.40*	2	.0355	4.55*	2	.0365	5.37*
Square uniqueness	2	.0395	1.64	2	.0055	<1.00	2	.0005	<1.00
Residual	42	.0241	—	18	.0078	—	42	.0068	—
Total	71	.0419	—	35	.0174	—	71	.0085	—

* $P < .05$.

rotated in a vertical plane around an axis at the center of *S*'s body to three body positions: L, E, and R. The aesthesiometers were attached to a platform mounted on the box in a fixed position relative to *S*'s body. Twelve *Ss* were tested. In Cond. 3 (Unsupported Head and Trunk Tilt) *S* stood and with head and trunk straight assumed three positions: head and trunk L, E, and R. A scale behind *S* assured that he maintained the required position. The aesthesiometers were mounted on a platform which was tilted in the same direction and amount as *S*'s head and trunk. Twenty-four *Ss* were tested.

Results.—The analysis of variance for Cond. 1 head tilt (Table 1), shows that the *F* test for tilt is significant. In general the left palm is more sensitive than the right; however the relative differences between conditions is in accord with the hypothesis. The relative sensitivity of the left palm is greatest for head tilt right (−2.00 mm.), less for head erect (−1.54 mm.), and least for head tilt left (−0.25 mm.) (Table 2).

The analysis of variance for Cond. 2, supported body tilt (Table 1), shows that the *F* test for tilt is significant. For all conditions the right palm is more sensitive than the left palm; however, the relative differences between the conditions are in accord with the hypothesis. The relative sensitivity of the right palm is greatest for head tilt left (+4.33 mm.), less for head erect (+3.08 mm.), and least for head tilt right (+1.34 mm.) (Table 2).

The analysis of variance for Cond. 3, unsupported head and trunk tilt (Table 1), shows that the *F* test for tilt is significant. The differences between the conditions are in accord

TABLE 2

EXPERIMENT 1: MEAN TWO-POINT THRESHOLDS (MILLIMETERS) FOR RIGHT AND LEFT PALMS

Tilt Condition	*N*	20° Left			Erect			20° Right		
		Left Palm	Right Palm	Diff. LP-RP	Left Palm	Right Palm	Diff. LP-RP	Left Palm	Right Palm	Diff. LP-RP
1. Head tilt	24	8.83	9.08	−0.25	8.92	10.46	−1.54	8.71	10.71	−2.00
2. Supported body tilt	12	11.16	6.83	+4.33	10.75	7.67	+3.08	9.92	8.58	+1.34
3. Unsupported body tilt	24	8.58	8.33	+0.25	9.17	9.88	−0.71	8.25	9.58	−1.33

with the hypothesis. The relative sensitivity of the right palm is greatest for head tilt left (+0.25 mm.), less for head erect (−0.71 mm.), and least for right head tilt (−1.33 mm.) (Table 2).

Experiment II

Generalizing from earlier findings showing a concomitant reorganization in visual space with a redistribution of visual sensitivity, the following hypothesis was formulated: tilt left (right) shifts the apparent median plane to the right (left). The same sets of postural conditions as in Exp. 1 are used here: head tilt, supported body tilt, and unsupported head and trunk tilt.

Method.—Each *S* made a tactual-kinaesthetic adjustment of a perpendicular rod to a position which appeared straight ahead of the body (in the apparent median plane). Independent groups of *Ss* were used for each set of postural conditions. As in Exp. I, each set of conditions consisted of three positions: tilt 20° left (L), erect (E), and tilt 20° right (R). For each of these positions two starting positions of the rod were used: 30 cm. left and 30 cm. right of the center of the body (the objective median plane). Thus there was a total of six trials which were arranged in a 6 × 6 replicated latin square design.

The perpendicular rod could be moved left and right in a horizontal track 6 in. in front of *S*. Location of the rod is measured in centimeters with respect to the objective median plane. With *S*'s eyes closed, his fingers are placed on the perpendicular rod. Instructions are to move the rod to the left or right until it is located straight ahead of the body.

The position of the apparent median plane in terms of deviations from the objective median plane is utilized. A plus deviation indicates a shift of the apparent median plane to the right and a minus deviation indicates a shift to the left.

Condition 1 (Head Tilt) was the same as in Exp. I, with 18 *Ss*. Condition 2 (Supported Body Tilt) was the same as in Exp. I, with 24 *Ss*. Condition 3 (Unsupported Head and Trunk Tilt) was also the same as in Exp. I, with 24 *Ss*.

Results.—The analysis of variance for Cond. 1, head tilt, shows that tilt is a significant source of variance (Table 3). The differences between conditions are in accord with the hypothesis. For head erect the apparent median plane is +0.06 cm. from the objective median plane. With left tilt, the apparent median plane shifts relatively to the right (+0.82 cm.), and with right tilt, the

TABLE 3

Experiment II: Analyses of Variance for Location of Apparent Straight Ahead

Source	1. Head Tilt			2. Supported Body Tilt			3. Unsupported Body Tilt		
	df	*MS*	*F*	*df*	*MS*	*F*	*df*	*MS*	*F*
Ss	17	8.815	1.08	23	21.279	1.06	23	15.457	1.01
Sequence (S)	5	10.294	1.25	5	25.707	1.28	5	15.955	1.04
Ss within S	12	8.199	—	18	20.049	—	18	15.318	—
Order	5	2.694	<1.00	5	26.601	2.32	5	20.864	1.97
Conditions	5	55.772	14.99*	5	175.887	15.33*	5	200.236	18.89*
Tilt (T)	2	12.380	3.33*	2	69.797	6.08*	2	98.461	9.29*
Starting position (SP)	1	250.560	67.33*	1	725.400	63.24*	1	641.356	60.51*
T × SP	2	1.770	<1.00	2	7.219	<1.00	2	81.452	7.68*
Square uniqueness	20	7.938	2.13	20	15.522	1.35	20	13.366	1.26
Residual	60	3.721	—	90	11.471	—	90	10.600	—
Total	107	7.703	—	143	19.892	—	143	18.757	—

* $P < .05$.

TABLE 4

EXPERIMENT II: MEAN LOCATION OF APPARENT STRAIGHT AHEAD (CENTIMETERS)

Tilt Condition	N	20° Left		Erect		20° Right	
		App.	Diff. Phys.-App.	App.	Diff. Phys.-App.	App.	Diff. Phys.-App.
1. Head	18	49.18	+0.82	49.94	+0.06	50.34	−0.34
2. Supported body	24	47.51	+2.49	48.57	+1.43	49.92	+0.08
3. Unsupported body	24	46.98	+3.02	48.33	+1.67	50.43	−0.43

Note.—Physical straight ahead = 50; less than 50 is right and greater than 50 is left of physical straight ahead.

apparent median plane shifts relatively to the left (−0.34 cm.) (Table 4).

The analysis of variance for Cond. 2, supported body tilt, shows that tilt is a significant source of variance (Table 3). Even though for all conditions the apparent median plane is located to the right of the objective median plane, the relative differences between these conditions are in accord with the hypothesis. For body erect the apparent straight ahead is +1.43 cm. from the objective median plane. With left tilt, the apparent straight ahead shifts relatively to the right (+2.49 cm.), and with right tilt, the apparent straight ahead shifts relatively to the left (0.08 cm.) (Table 4).

The analysis of variance for Cond. 3, unsupported head and trunk tilt, shows that tilt is a significant source of variance (Table 3). Differences between conditions are in accord with the hypothesis. For head and trunk erect the mean location of the apparent straight ahead is +1.67 cm. from the objective median plane. With left tilt the apparent straight ahead shifts relatively to the right (+3.02 cm.), and with right tilt the apparent straight ahead shifts relatively to the left (−0.43 cm.) (Table 4).

DISCUSSION

With respect to the first aim of these studies, the findings for normal adults are in accord with previous findings for cerebellar patients (Halpern & Kugelmas, 1957): the distribution of tactual sensitivity depends on the postural status of the organism. Further, we can say that the direction of postural change, body tilt left or right, has a systematic effect on the distribution of tactual sensitivity, viz., sensitivity increases right or left, respectively. On this basis we may propose that a comprehensive explanation of tactual sensitivity must take into account postural factors as well as the tactual afferent system.

In regard to our second aim, the present findings permit generalization of the previously found relationship in the visual modality to the tactual modality: there are concomitant changes in the distribution of sensitivity and the organization of space depending on postural changes.

A variety of studies conducted by other investigators may be adduced in support of the generalization offered here that postural changes also affect the distribution of tactual sensitivity and the tactual-kinaesthetic organization space. Concerning sensitivity, both temporary and enduring changes due to postural factors have been reported. As to temporary changes, it has been reported (Frank, 1957) that when the hand and fingers are resting inactive the third finger is the most sensitive but when the hand and fingers are actively searching the forefinger is the most sensitive. Enduring changes, due to amputation and brain injury have been reported respectively by Teuber, Krieger,

and Bender (1949); Haber (1955); and Halpern and Kugelmas (1957). Teuber and Haber have found that the sensitivity on all remaining parts of an amputated arm is proportionally greater than on the homologous parts of the amputee's intact arm. Halpern and Kugelmas report that cerebellar patients who manifest balanced asymmetrical posture to one side show an increase of tactual sensitivity on the opposite side of the body. These changes in tactual sensitivity—both transient and enduring, point to the role of postural factors in the distribution of sensitivity: a change in sensitivity with a shift from inactivity to activity of the hand; an increase in sensitivity in the stump of the amputee's arm when it acquires the function of operating as a manipulative organ; finally, the change in tactual sensitivity with a shift from a symmetrically balanced to an asymmetrically balanced stance.

Temporary and enduring changes have also been reported in the tactual-kinaesthetic organization of space. In regard to temporary changes, Werner, Wapner, and Bruell (1953) have shown that head turning systematically shifts the direction of pointing when subjects are asked to point to the straight-ahead.[2] Goldstein (1936) has reported enduring changes for cerebellar patients who adopt abnormal postures: they shift the locus of the middle of their bodies depending on the side of lesion. Halpern (1951) has also found that patients with cerebellar and vestibular disturbance shift the location of the middle of their bodies.

The work of others just reviewed, an earlier set of studies in the visual modality, and the experiments presented here, all point to the fact that with changes in postural status there are changes *both* in the distribution of sensitivity and in the organization of space. We may, however, go one step further. Our previous studies in vision and the present studies in the tactual modality point to systematic directional shifts in the spatial distribution of sensitivity and in the organization of space; that is, with an increase of sensitivity to one side, there is a shift of the apparent median plane to that side. It seems therefore plausible on this basis to postulate a bond between these two shifts—in the distribution of sensitivity and the organization of space—occurring with changes in postural status. As is well known, some theorists such as Goldstein (1939, pp. 53–55) would view this relationship in terms of its value for the adaptation of the organism. Our eventual aim, however, is to account for this bond between sensitivity and space organization in terms of a mechanism which permits the reciprocal interaction of sensory and postural factors.

[2] We may refer here also to related studies by Morant (1952) which demonstrate systematic shifts in the tactual-kinaesthetic median plane depending on labyrinthine stimulation due to rotation.

Summary

Experiments have been reported which show the effect of postural tilts left and right on (*a*) the distribution of tactual sensitivity as measured by the two-point threshold of the palms, and on (*b*) the organization of tactual-kinaesthetic space as measured by the location of the tactual-kinaesthetic straight-ahead. Tilt left increases sensitivity of the right palm compared to the left palm, and also shifts the straight-ahead to the right. Tilt right increases sensitivity of the left palm compared to the right palm, and also shifts the straight-ahead to the left. These findings indicate the necessity of considering postural factors as well as the tactual afferent system in a comprehensive explanation of sensitivity. Further, these findings are viewed as extending the previously found relationship between the distribution of visual sensitivity and the organization of visual space depending on postural factors: that is, for both the visual and the tactual modalities there are systematic directional shifts depending on postural factors. This evidence makes it plausible to postulate an inner relationship between these two types of shifts —distribution of sensitivity and organization of space.

References

Boring, E. G. *Sensation and perception in the history of experimental psychology.* New York: Appleton-Century, 1942.

BRUELL, J. Visual egocentric localization, an experimental study. Unpublished doctoral dissertation, Clark University, 1953.

FRANK, L. K. Tactile communication, *Genet. psychol. Monogr.*, 1957, **56**, 209–256.

GIBSON, E. J. Improvements in perceptual judgments as a function of controlled practice or training. *Psychol. Bull.*, 1953, **50**, 401–432.

GOLDSTEIN, K. The function of the cerebellum from a clinical standpoint. *J. nerv. ment. Dis.*, 1936, **83**, 1–12.

GOLDSTEIN, K. *The organism.* New York: American Book., 1939.

HABER, W. B. Effects of loss of limb on sensory functions. *J. Psychol.*, 1955, **40**, 115–123.

HALPERN, L. The syndrome of sensory-motor induction in combined cerebellar and labyrinthine injury. *J. nerv. ment. Dis.*, 1951, **114**, 137–161.

HALPERN, L., & KUGELMAS, S. The variability of tactile two-point discrimination in the sensory-motor induction syndrome. *Acta psychol., Amst.*, 1957, **13**, 162–172.

JENKINS, W. L. Somesthesis. In S. S. Stevens (Ed), *Handbook of experimental psychology.* New York: Wiley, 1951, Pp. 1172–1190.

McFARLAND, J. H. The effect of asymmetrical eye torsion on perceptual clarity. Unpublished master's thesis, Clark University, 1957.

McFARLAND, J. H., WERNER, H., & WAPNER, S. The effect of muscular involvement on sensitivity: Asymmetrical convergence on the distribution of visual sensitivity. *Amer. J. Psychol.*, 1960, **73**, 523–534.

MORANT, R. B. Factors influencing the apparent median plane under conditions of labyrinthian stimulation. Unpublished doctoral dissertation, Clark University, 1952.

TEUBER, H.-L., KRIEGER, H. P., & BENDER, M. B. Reorganization of sensory function in amputation stumps: Two-point discrimination. *Fed. Proc.*, 1949, **8**, 156.

WEBER, E. H. Der Tastsinn und das Bemeingefuhl. In R. Wagner (Ed), *Handwoerterbuch der Physiologie III.* Braunschweig: Vieweg, 1846, Pp. 481–588.

WERNER, H., WAPNER, S., & BRUELL, J. Experiments on sensory-tonic field theory of perception: VI. Effect of position of head, eyes, and of object on position of the apparent median plane. *J. exp. Psychol.*, 1953, **46**, 293–299.

(Received November 16, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 2, 155–159

ACQUISITION AND TRANSFER IN PARTIAL REINFORCEMENT

E. J. CAPALDI AND MONTE G. SENKO[1]

University of Texas

According to the aftereffects hypothesis, reinforcement (or nonreinforcement) on a particular trial serves to give rise to distinctive stimuli which may persist until the following trial and become conditioned to responses. Although the hypothesis was originally formulated to deal with extinction (Sheffield, 1949), it may also be employed in connection with acquisition and transfer phenomena. Experimental results relevant to these latter aspects of the hypothesis have recently become available (Bloom & Capaldi, 1961). However, these data are quite limited; neither are they as critically related to the hypothesis as seems desirable, nor do they exclude alternative interpretations with a degree of sufficiency which seems necessary. The major purpose of this investigation was to test specific deductions available from an aftereffects hypothesis with regard to acquisition and transfer using regular patterns of reinforcement. There were certain lesser purposes as well, which will be discussed as occasion arises.

Consider a reinforcement pattern in which two consecutive nonreinforced trials are consistently followed by a reinforced trial (regular 33%). In this case the aftereffects of nonreinforcement are equally often followed by reinforcement and nonreinforcement, while the aftereffects of reinforcement are followed only by nonreinforcement. Prediction from the hypothesis in this instance is clear. It is expected that completely appropriate pattern running, such as develops in the single alternation situation (SA) will not occur (Bloom & Capaldi, 1961; Capaldi, 1958; Tyler, Wortz, & Bitterman, 1953). Rather, the hypothesis predicts that the rat will run appropriately (relatively rapidly) on the reinforced trial (R), appropriately (relatively slowly) on the nonreinforced trial following the reinforced one (N_1), and inappropriately (relatively rapidly) on the succeeding nonreinforced trial (N_2). It is further expected that running times on Trials R and N_2 should be practically identical.

Assuming that the rat responds in the manner outlined above, further predictions regarding transfer are available from the hypothesis. As indicated, the hypothesis holds that during acquisition in the regular 33% situation *S* will form the habits of running rapidly following nonreinforcement on the one hand, and slowly following reinforcement on the other. But, according to the hypothesis, these are precisely the habits which are formed in the SA situation. Therefore, upon transfer, under the conditions outlined above, it is expected that immediate, consistent, and completely appropriate SA pattern running will occur.

METHOD

Subjects.—The *S*s, approximately 95 days old at the beginning of the experiment, were 31 male Wistar strain rats purchased from Barlan Industries, Cumberland, Indiana.

Apparatus.—The apparatus was a straight-alley runway enclosed by walls, with a battery of five goal boxes mounted on runners.

[1] Now at the University of Wisconsin.

Each goal box could be positioned behind a one-way, gray, swinging door which separated goal box from alley. Goal boxes not in line with the runway were neither visible nor accessible to *S* from the alley. Placing *S* down at the start point of the alley depressed an $8 \times 3\frac{1}{8}$ in. platform supported by a microswitch and simultaneously activated a timer. Upon touching the $3\frac{3}{4} \times 6$ in. swinging door at the alley's end, contacts were broken and the timer, which measured in .001 sec., was stopped. The elapsed time was recorded.

The dimensions of the alley, which was painted gray, were $66 \times 3\frac{3}{4} \times 6$ in. The colors of the goal boxes from left to right were gray, gray, white, black, and white. Reinforcement occurred in one gray or white box and nonreinforcement in the other gray or white box to control for the presence of food particles on nonreinforced trials. The inside dimensions of each goal box were $11\frac{1}{2} \times 4\frac{1}{4} \times 6$ in. A hinged, hardware-cloth lid covered the entire battery of goal boxes. Illumination was provided by a fluorescent lamp placed 5 ft. above the starting point.

Procedure.—On Days 1 through 5 *S*s were handled and fed wet mash 1 hr/day. On Day 6 each *S* was given six preliminary trials, three from 1 ft. in front of the goal box and three from a point about halfway in the alley. On this day the swinging door was raised and food was present in the goal box. On Day 7 each *S* received six runs from the start point of the alley with swinging door down and food in the goal box. Throughout all phases of the experiment *S* was confined to the goal box for 20 sec. following each run, the intertrial interval was 20 sec., and *S* was fed for 1 hr. in the home cage immediately following the daily runs.

Experimental training was begun on Day 8. All *S*s were given 12 trials per day and 33% reinforcement. For Group IR-G ($N = 8$) reinforcement was random and both reinforcement and nonreinforcement occurred in a gray goal box. Group R-G ($N = 8$) was comparable in all respects save that reinforcement was regular—the initial trial of the day being nonreinforced, the third reinforced, and so on. Finally, Group R-WBW ($N = 15$) was likewise regularly reinforced; however, on trials preceding reinforcement *S* found a black goal box and on trials preceding nonreinforcement, a white goal box. The *S* was thus both nonreinforced (Trials 1, 4, 7, and 10) and reinforced (Trials 3, 6, 9, and 12) in a white goal box.

On Day 15 all three groups were transferred to an SA pattern. Both reinforced and nonreinforced trials occurred as before in gray goal boxes. Group R-G received 7 days of SA training, Group IR-G received 9 days of such training. Group R-WBW was trained as before on the regular 33% pattern for an additional 7 days, i.e., a total of 21 days. The intertrial interval, etc., were as before.

Results

In Fig. 1 are presented the log medians of median running times of Group R-WBW on Trials N_1, N_2, and R, on each of the 21 days of acquisition training. On the initial days *S*s ran relatively slowly on R trials and increasingly more rapidly on Trials N_2 and N_1, respectively. The Day 1 and Day 2 differences (analysis using log medians) were significant ($F = 10.43$ and 8.36, $df = 2/28$ each case, $P < .01$ each case). The tendency for slow running on N_1 trials began on Day 6 and increased consistently thereafter until Day 11, when it seems to have reached full strength. As can be seen, Trial N_2 and R running times are practically identical from about Day 6 onward. While the differences shown in Fig. 1 on Day 6 appear small, they are nevertheless significant ($F = 8.95$, $df = 2/28$, $P < .01$). A separate analysis of N_2 and R times failed to reveal significant differences ($F < 1$), thus differences in the overall analysis must be due to slower running on N_1 trials. Actually, on Day 6, 13 of the

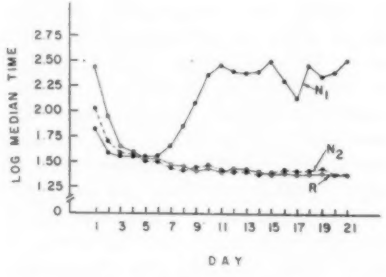

Fig. 1. Log median of median running times of Group R-WBW for each of the acquisition days.

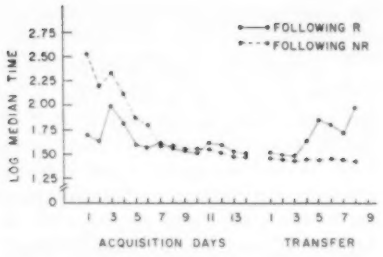

FIG. 2. Log median of median running times of Group IR-G for each of the acquisition and transfer days.

15 *S*s ran more slowly on N_1 trials than on either N_2 or R trials.

Since, later in training, *S*s ran equally rapidly on both N_2 and R trials it appears that the presence of the white goal box on all trials preceding nonreinforcement did not serve as a particularly effective cue from the point of view of pattern running. If it had, that is, if color of goal box gave rise to distinctive aftereffects, as apparently does food or its absence, then N_2 running times presumably should have been slower than those on R trials.

In Fig. 2 the log median of median running times of Group IR-G on trials following reinforcement and on trials following nonreinforcement on each of the acquisition and transfer days are presented. On each of the initial 6 days of acquisition training, *S*s ran more rapidly on trials following reinforcement. These daily differences, except for those on Day 5, are significant from beyond the .05 to beyond the .01 level ($F = 14.30$, 19.02, 6.33, 15.30, 4.20, and 7.13 for Days 1 through 6 respectively; $df = 1/7$ in each case). On Days 7 through 10, running speeds are about equal on both types of trials; however, from Day 11 onward *S*s show a slight tendency to run more slowly on trials following reinforcement. The log of the sums of the median running speeds for each *S* on trials following reinforcement and on trials following nonreinforcement on Days 11 through 14 were analyzed. The differences very closely approach but do not reach significance at the .05 level ($F = 5.37$; needed at the .05 level for 1/7 *df*, an *F* of 5.59). The numbers of transitions of each type on the initial 10 days of training were R-R, 12; R-N, 24; N-R, 28; and N-N, 46. Thus, the percentage of reinforced trials followed by a reinforced trial is 33.3 and the percentage of nonreinforced trials followed by a reinforced trial is 37.8. It seems fair, if not necessary, to conclude that the differential running was due to the slight differences in percentage indicated above since, when nonreinforced trials are followed by an exactly equal number of reinforced and nonreinforced trials, identical running times on N_2 and R trials result. As will be seen, this was true not only for Group R-WBW but for Group R-G as well. That such small differences may result in differential running seems surprising.

The tendency to run more slowly following reinforcement persisted into the transfer stage. Analysis of log median differences on the initial day of SA training yielded an *F* of 4.75 which, for 1/7 *df*, is not significant. Despite the fact that under the SA pattern R trials are consistently followed by N trials, the already slightly developed tendency to run slowly following R trials did not increase in strength until Day 4 of training. By this day the differences reached significance ($F = 6.49$, $df = 1/7$, $P < .05$).

In Fig. 3 are presented the log median of median running times of Group R-G during acquisition and transfer. In the acquisition stage these times are presented on the basis of R, N_1, and N_2 trials, and in the transfer stage, on the basis of trials

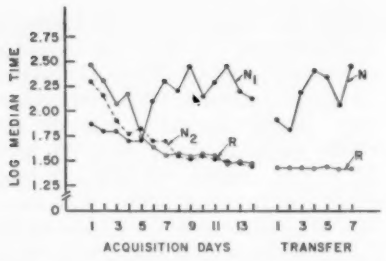

FIG. 3. Log median of median running times of Group R-G for each of the acquisition and transfer days.

following reinforcement and trials following nonreinforcement. On each of the initial 4 days of acquisition, as before, running times were most rapid on N_1 trials, less rapid on N_2 trials, and least rapid on R trials. Only the differences on the initial day were significant ($F = 22.16$, $df = 2/14$, $P < .01$). Running times on N_2 and R trials are practically identical from Day 8 onward, and the tendency for slow running began on Day 6 and appears to have reached full strength on Day 9. As in the case of Group R-WBW, significantly slower running on N_1 trials appeared initially on Day 6 ($F = 4.93$, $df = 2/14$, $P < .05$).

When shifted to the SA pattern, *Ss* showed a tendency to run more slowly on nonreinforced trials, i.e., on trials following reinforcement. However, Fig. 3 does not show an important aspect of these data. The least rapid of the initial day's reinforced trial running times was shorter than the shortest of that day's nonreinforced trial running times. In other words, on the initial day, a saw-toothed, SA pattern running curve was obtained when the data were plotted on a trial by trial basis. An analysis of the log median times on reinforced vs. nonreinforced trials on the initial day showed them to be significant ($F = 22.42$, $df = 1/7$, $P < .01$).

It should be mentioned as an examination of Fig. 1 and Fig. 3 will indicate, that in general Group R-WBW ran more rapidly than Group R-G. Considering daily performance, the only exceptions to this tendency on N_1 trials occurred on Days 10, 11, 13, and 14; and on R trials, on Day 1. No exceptions occurred considering N_2 trials. Daily differences between the groups were tested for each type of trial for each of the initial 8 days. In each case, $df = 1/21$ and, with a single exception, differences reported were significant beyond the .05 level. Group R-WBW ran significantly more rapidly than Group R-G on R trials on Day 3 ($F = 5.72$), Day 4 ($F = 7.11$), and Day 5 ($F = 4.96$); on N_1 trials on Day 4 ($F = 11.46$, $P < .01$); and on N_2 trials on Day 3 ($F = 4.36$), Day 4 ($F = 4.48$), Day 5 ($F = 6.33$), Day 6 ($F = 6.17$), and Day 7 ($F = 7.24$). Since the greatest effect seems to have been on N_2 and R trials, i.e., on trials following nonreinforcement, rather than on N_1 trials, which followed reinforcement, it seems that Group R-G was more susceptible to relatively slow running following nonreinforcement.

DISCUSSION

The predictions of major concern have been consistently borne out by the data and need little comment. What does appear to be in need of comment, however, is the relevance of the present results for an alternative hypothesis which was the original source of and has been consistently applied to pattern learning experiments of the present type. In addition, this hypothesis is compatible with the results of most investigations which have employed irregular patterns of reinforcement.

According to the serial patterning version of the discrimination hypothesis,

partial reinforcement cannot be understood on a trial by trial basis but requires, rather, a gestalt-like level of analysis which considers that the rat responds to the entire series of events, i.e., learns a pattern per se (Capaldi, 1958; Tyler et al., 1953; Wodinsky & Bitterman, 1959). With reference to the initial proposition, that a trial by trial analysis is theoretically improper, the experimental results would seem to speak for themselves. That the rat learns a pattern per se, the second of the proposals entertained by the hypothesis, seems quite unlikely from these results. Actually, two somewhat distinct criticisms are implied here. First, a trial by trial aftereffects analysis is able to explain that which the serial patterning hypothesis purports to explain but in fact merely accepts without explanation, e.g., pattern running in the SA situation. Moreover, in specific situations, the aftereffects hypothesis yields predictions utterly unavailable to the patterning hypothesis. As far as can be determined, a patterning view cannot be made to indicate whether the rat will respond completely appropriately in the presence of a regular 33% pattern and, if not, in what specific respects response will be inappropriate. Nor do specific predictions regarding transfer seem available from the patterning hypothesis.

The successes and failures of a discrimination hypothesis do not appear difficult to understand. That hypothesis, in common with a pure form of the aftereffects hypothesis, holds that resistance to extinction and amount of stimulus change are inversely related. Apparently, a simple examination of reinforcement patterns, despite a lack of knowledge of specific mechanisms, is adequate in most instances to determine which of two or more patterns most differs from the extinction situation, i.e., zero reinforcement. While this procedure has encountered some success as regards extinction, it appears less adequate in the case of transfer and entirely inadequate with reference to acquisition. As the present results have shown, a knowledge of specific mechanisms is required to achieve prediction in these latter instances.

Summary

Three groups of rats were trained to traverse a runway under conditions of 33% reinforcement. Two of the groups were regularly reinforced. One of the regular groups was both reinforced and nonreinforced in a gray goal box. The other regular group was reinforced in a white goal box and nonreinforced in a white and in a black goal box. The third group was irregularly reinforced in a gray goal box. As predicted from an aftereffects hypothesis, the regular groups ran slowly following reinforcement and rapidly on the succeeding two trials. The regular gray group was transferred to a single alternation (SA) pattern of reinforcement. Immediate and consistent SA pattern running occurred. Color of goal box did not appear to be a factor in giving rise to distinctive aftereffects.

It was suggested that certain of the results could not be accounted for by the serial patterning version of the discrimination hypothesis.

References

Bloom, J., & Capaldi, E. J. The behavior of rats in relation to complex patterns of partial reinforcement. *J. comp. physiol. Psychol.*, 1961, **54**, 261–265.

Capaldi, E. J. The effect of different amounts of training on the resistance to extinction of different patterns of partially reinforced responses. *J. comp. physiol. Psychol.*, 1958, **51**, 367–371.

Sheffield, V. F. Extinction as a function of partial reinforcement and distribution of practice. *J. exp. Psychol.*, 1949, **39**, 511–526.

Tyler, D. W., Wortz, E. C., & Bitterman, M. E. The effect of random and alternating partial reinforcement on resistance to extinction in the rat. *Amer. J. Psychol.*, 1953, **66**, 57–65.

Wodinsky, J., & Bitterman, M. E. Partial reinforcement in the fish. *Amer. J. Psychol.*, 1959, **72**, 184–199.

(Received November 17, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 2, 160–166

AMOUNT, DELAY, AND POSITION OF DELAY OF REINFORCEMENT AS PARAMETERS OF RUNWAY PERFORMANCE [1]

EUGENE R. WIST [2]

University of Missouri

The introduction of delay at some point in a sequence of behavior can exercise its effect in two places, on the responses preceding the point in the response chain where delay is introduced and on the responses following this point.

The effect of delay on the responses preceding the point of delay is reported as a gradual decrement in response strength over a series of trials to a new asymptote (Hull, 1952; Spence, 1956). Furthermore a gradient of response strength, the "goal gradient," is supposed to exist among the members of the response chain terminating in reinforcement (Hull, 1934). Thus the new asymptote after delay introduction should be some function of the point in the chain at which delay is introduced.

The effect of delay on the responses following the point of delay has been reported both as incremental (e.g., Bernstein, 1954) and decremental (e.g., Collier & Marx, 1957). Again if the effect, whatever it is, is proportional to the predelay response strength, then its magnitude would be some function of the point in the chain at which delay is introduced.

The present experiment attempts to make a direct examination of these conjectures by means of two, independent, presumed manipulators of response strength. First, the point in the response chain at which delay is introduced was varied. If a within-chain gradient exists, this should result in variations in response strength at different points in the chain. Second, the magnitude of reinforcement was varied to produce two degrees of response strength. Combinations of these two variables and different delay durations made possible a comparison of the effects of the two means of producing variation in response strength on pre- and post-delay running gradients. Furthermore, quantitative measurement and control of consummatory behavior made possible a determination of some of the relationships between end-box events and running gradients.

METHOD

Experimental design.—Two reinforcement concentrations (8% and 32% sucrose solutions), two delay intervals (3 and 45 sec.), and three positions of delay (2, 6, and 10 ft. from the goal box) were factorially combined. Two additional 8% and 32% control groups were run without delay.

Subjects.—The *Ss* were 98 female albino rats with a mean age of approximately 4 mo.;

[1] This experiment represents a portion of a dissertation submitted to the Psychology Department of the University of Missouri in partial fulfillment of the requirements for the PhD degree. The author wishes to express his indebtedness to George Collier, and to Melvin H. Marx, his advisor, for their advice and encouragement. A report on a portion of these data was made at the 1960 Midwestern Psychological Association meetings in St. Louis.

This investigation was supported in part by research grants, M-817 and M-3328, from the National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland.

[2] Now at the United States Army Medical Research Laboratory, Fort Knox, Kentucky.

53 *S*s were of the Holtzman strain, 45 were from the colony of the Department of Psychology, University of Missouri. Two rats were discarded for failure to run or drink, 1 because of illness, 4 because of injury in the apparatus, and 7 to equalize *N*s at 6 per cell. All *S*s were experimentally naive, and were assigned at random to groups.

Apparatus.—The apparatus was an 18-ft. straight runway housed in a light-tight, sound-deadened hull. The start box, stem, and goal box were 1½ ft., 14 ft., and 2½ ft. long, respectively. Two guillotine doors separated these boxes from the stem. The delay box consisted of a pair of movable guillotine doors separated by 2 ft. and mounted on a frame. The rear door was operated by *S*'s entry into the chamber, the transparent front door by means of a Hunter decade timer. Similarly the end-box door was operated by *S*'s entry.

The runway floor consisted of nine 2-ft. sections of angle aluminum. The start-box door was located 6 in. short of the end of the first section. The goal-box door was located 6 in. short of the beginning of the last section. Similarly, the position of the delay box was arranged so that the front and rear doors were located 6 in. before the ends of the sections they isolated.

The runway was automatic. The start-box door opened and the first clock was started 3 sec. after the start-box lid was lowered. When *S*s reached the end of the first 2-ft. section, the first clock stopped and the second started. When the third section was reached, the second clock stopped and the third started, etc.

Solutions were delivered by two 50-ml. burettes to each of which a drinking tube with a 4-mm. opening had been annexed. These tubes were ⅛ in. behind a ½-in. square opening in a panel located 6 in. from the end of the goal box. A 7-w. light 5 in. behind this panel illuminated the goal box and runway. Contact with the tube activated an electronic relay. Number of licks, duration of tube contact, and volume consumed were recorded.

The sucrose solutions were mixtures of commercial sugar and distilled water made 24 hr. before their use and maintained at room temperature.

A four compartment box simulating the goal box of the runway was used as a drinker trainer.

Pretraining.—Twenty-eight days before the first runway trial, *S*s were placed on a 22-hr. food deprivation schedule. On Day 13 the feeding period was reduced to 1 hr. per day and all *S*s were placed on a table top in groups of 6 and allowed to explore for 15 min. during which time *E* picked up and handled each *S* for two 1-min. periods. On Day 17 each *S* received 5 min. of drinking in the drinking trainer. Half of the *S*s received an 8%, and half a 32% sucrose solution. On Day 18, this period was reduced to 3 min. On Days 19 through 25, *S*s received 10 min. of exploration on top of the runway followed by a 10-min. adaptation period in the runway. The floor of the runway was strewn with .45 mg. Noyes food pellets. On Day 26 *S*s were again placed in the drinking trainer for a 3-min. period of drinking. On Days 27 and 28 of deprivation, *S*s were individually placed in the goal box until a total of 15 sec. of tube contact was recorded.

Runway trials.—The *S*s were given one trial per day for 52 days. For the first 26 days, no delay was administered, but the delay-box doors were in operation at the appropriate position in the runway allowing *S*s to adapt to the auditory, visual, and tactual cues resulting from their operation. For the last 26 trials delay was administered on each trial at the appropriate positions. The *S* was allowed a 20-sec. drinking period in the goal box after which it was removed and returned to its home cage.

Results and Discussion

Three 5-trial blocks were analyzed. The first block, Trials 22–26, consisted of the last 5 trials before the introduction of delay. The second block, Trials 27–31, consisted of the first 5 delay trials. The last block, Trials 48–52, consisted of the last 5 trials. Analyses of variance (Lindquist Type I and Type VI) were performed on both runway and consummatory behavior measures. Table 1 presents the major significant results of these analyses. The complete analyses may be found elsewhere (Wist, 1960). Alexander (1946) tests of trend were performed on the 15 trials preceding the delay introduction for each concentration group independently. None of the *F*s was significant, suggesting that running speed was asymptotic at the time of delay introduction.

TABLE 1

MIXED ANALYSES OF VARIANCE PERFORMED ON RUNNING AND CONSUMMATORY BEHAVIOR MEASURES

Measure	Source*	*F*	*df*
Starting speed, Start box (Fig. 2A)	C	35.31**	1/70
	C × B	7.68**	1/140
	D × B	18.32**	2/140
Running speeds, Predelay box segment (Fig. 2B)	C	21.98**	1/70
	D × B	26.91**	2/140
	P (Trials 22–26)	22.03	1/60
Starting speeds, Delay box (Fig. 2C)	C	4.58*	1/70
	B	28.57**	2/140
	Delay	30.22**	1/70
	B	32.03**	1/70
	D × B	6.21**	1/70
Running speeds, Postdelay box segment (Fig. 2D)	C	25.13**	1/70
	B	65.16*	2/140
	D (Trials 27–31)	17.27**	1/60
	B	78.35**	2/140
	D × B	14.71**	2/140
	P (Trials 22–26)	5.71*	1/60
Number of licks (Fig. 3A)	C (Trials 22–26)	5.02*	1/60
Lick rate (Fig. 3D)	C (Trials 22–26)	9.29**	1/60
Ingestion rate (Fig. 3E)	C × B	3.63*	2/140

Note.—Only major significant results are included in this table.

* C = Concentration, D = Delay, B = Blocks, P = Position.

* $P < .05$.

** $P < .01$.

Five of the results of the present experiment are of special significance.

First, the "classical" gradient of increased speed in the direction of the goal was not obtained (Hull, 1943). Instead, there was an initial burst of speed followed by a slight decrement as the goal was approached, a finding typical of this runway (Knarr, 1959; Knarr & Collier, 1960). Figure 1 shows the running gradients for the control groups for the 5-trial blocks immediately preceding and following the introduction of delay for the experimental groups. The experimental group gradients were similar in form immediately before delay introduction (Trials 22–26). This figure also indicates that the operation of the goal-box doors without delay had no appreciable effect on running speed (Trials 27–31).

The discrepancy between the data of the present study and those which have provided evidence for the goal gradient (Hull, 1934) would seem to be due to the fact that external stimuli were minimal in the runway employed. It would seem that this goal gradient is obtained when runway stimuli can acquire control over units of the response chain or occasion competing behavior (Pereboom, 1958). Control over the chain by terminal reinforcing and discriminative stimuli produces running gradients similar to those of the present study (Collier, Knarr, & Marx, 1961).

The above failure to obtain a goal gradient is reflected in the lack of a differential effect of position of delay of reinforcement on running speeds. The two significant *F*s for position (Trials 22–26) in Table 1 reflect respectively the initial burst of speed and the terminal decline of speed before delay introduction (see Fig. 2B and D).

Second, delay resulted in decremental changes in performance (see Fig. 2). This finding is contrary to the implications of "frustration-drive" models (e.g., Amsel, 1958;

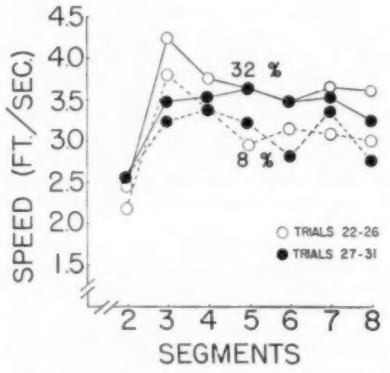

FIG. 1. Mean running speed for all segments over Trials 22–26 and Trials 27–31 for the control groups.

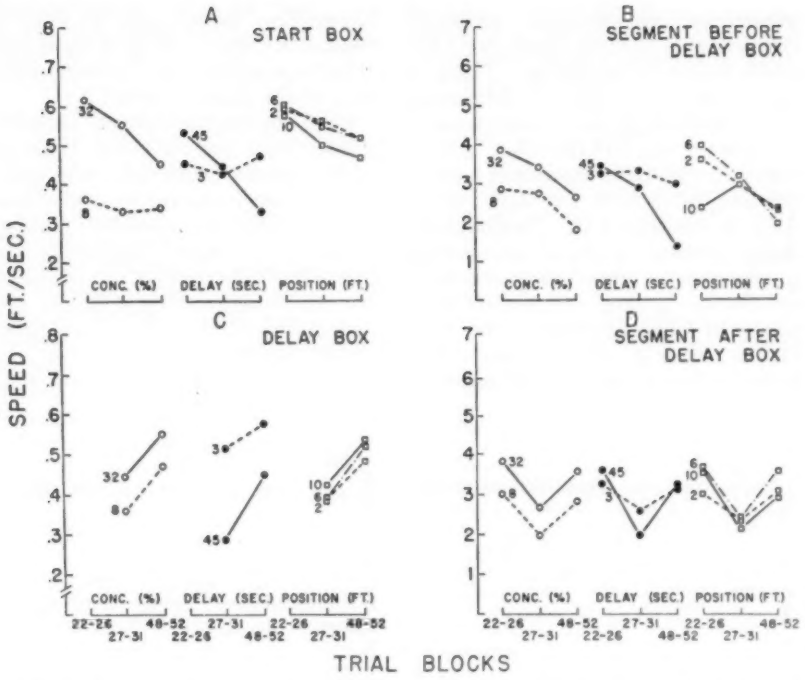

FIG. 2. Mean starting speeds for concentration, delay, and position groups from the start box (A) and delay box (C), and mean running speeds from the segment before the delay box (B) and the segment after the delay box (D) over trial blocks.

Brown & Farber, 1951). Since delay was not introduced until the second block of trials, there are only two points in Fig. 2C.

Third, delay had a differential effect on pre- and postdelay-box running speeds. Running speeds in the segments preceding the point of delay declined gradually to a new asymptote inversely proportional to delay duration. Running speeds in the segments following the delay box dropped immediately to a level proportional to the delay duration and then recovered gradually over trials to their original level (see Fig. 2B and D). These effects were significant (see Table 1, 2B, and 2D), and similar to those obtained by Collier and Marx (1957).

The effect on the responses preceding the point of delay is consistent with previous findings (Hull, 1952; Spence, 1956). According to the classical Hullian account delay weakens the habit strength (or reaction potential) of responses made more remote in time from the goal. According to this account the delay per se should not have affected postdelay-box running as it did, since the temporal relation of responses in this portion of the runway to the goal remained unchanged.

Fourth, delay and amount of reinforcement do not appear to be equivalent parameters of reinforcement. Vari-

ations in concentration produced differences in running speed along the whole length of the runway. Delay of reinforcement produced differences which were largely *local*. That is, maximum in the region of the delay. An exception to this latter statement was the starting speed from the start box of 32% *Ss* (see Fig. 2A). Furthermore, delay resulted in equal decrements performance for concentration groups, but greater decrements were produced for longer delays over trial blocks (see Fig. 2B and 2D). The latter finding is contrary to a response competition model such as that of Pereboom (1957). Both of these sets of findings are contrary to the implications of the classical model.

Fifth, the relation between consummatory and instrumental behavior in the present experiment was minimal. Three measures of consummatory behavior were recorded: number of licks, amount ingested, and duration of tube contact. In addition, two rate measures, rate of licking and rate of ingestion in terms of tube-contact duration were computed. In general, Fig. 3A-E indicates that

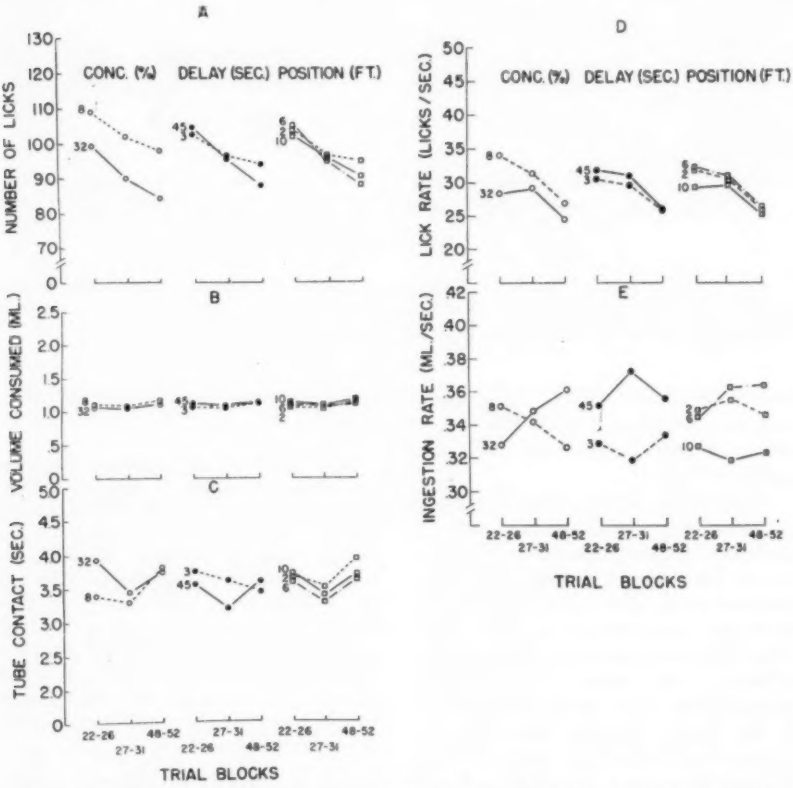

FIG. 3. Consummatory behavior measures for the three 5-trial blocks. (The lick rate and ingestion rate measures of 3D and 3E represent, respectively, the number of licks and amount ingested per second of tube contact.)

consummatory behavior was not greatly influenced by concentration, delay, or position of delay. Concentration, however, did influence initial lick number and rate and resulted in an interaction between concentration and blocks for ingestion rate (see Table 1).

A series of product-moment correlations was performed in which each of the nine dependent variables was correlated with every other one. This was done for the sums of each of the three 5-trial blocks upon which the analyses of variance were performed making a total of 108 *r*'s. All 72 *S*s entered into each correlation.

The most general finding was that while runway measures showed some relation to each other and consummatory measures similarly showed some interrelations, runway and consummatory measures were unrelated. Of the 60 possible correlations between runway and consummatory behavior only 3 reached significance at the .05 level. A complete summary of these relations is available elsewhere (Wist, 1960). This result supports the view that consummatory events have no necessary relation to instrumental behavior (Collier et al., 1961) and fails to support r_g models (e.g., Hull, 1952; Spence, 1956).

The above data, against the background of previous findings, suggests the following tentative formulation of the relationship between running gradients and variables affecting performance: "Running" occurs at rates determined by its consequences, such as the concentration of reinforcement, and it occurs everywhere at the same rate independent of its temporal or spatial distance from its climax. Factors such as fatigue or habituation may result in declines in rate over long distances. Factors such as obstacles and distracting stimuli may produce local disturbances in running, but do not manipulate its basic rate. Time and distance factors, important in the classical treatment of performance gradients do not readily account for the effects obtained.

Summary

When delay interrupts a response chain terminating in reinforcement, the effect on the within-chain gradient is presumably some function of the response strength of the link delayed. In the present study response strength was manipulated in two ways: (*a*) by varying the position of the delay in the chain; (*b*) by varying the amount of terminal reinforcement. Two reinforcement concentrations (8% and 32% sucrose solutions), two delay intervals (3 and 45 sec.), and three delay positions (2, 6, and 10 ft. from the goal box) were factorially combined.

Five major results were discussed: (*a*) The failure to obtain the goal gradient; (*b*) the decremental changes in performance due to delay; (*c*) the differential effect of delay on pre- and postdelay-box running speeds; (*d*) the nonequivalence of delay and amount of reinforcement as parameters of reinforcement, and (*e*) the minimal relation found between consummatory and instrumental behavior.

A tentative formulation of the relationship between running gradients and reinforcement parameters was suggested.

References

Alexander, H. W. A general test for trend. *Psychol. Bull.*, 1946, **43**, 533–557.

Amsel, A. The role of frustrative non-reward in noncontinuous reward situations. *Psychol. Bull.*, 1958, **55**, 102–119.

Bernstein, B. B. Extinction as a function of frustration. Unpublished doctoral dissertation, University of Missouri, 1954.

Brown, J. S., & Farber, I. E. Emotions conceptualized as intervening variables with suggestions toward a theory of frustration. *Psychol. Bull.*, 1951, **48**, 465–495.

Collier, G., Knarr, F. A., & Marx, M. H. Some relations between the intensive properties of the consummatory response and reinforcement. *J. exp. Psychol.*, 1961, **62**, 484–495.

Collier, G., & Marx, M. H. Frustration and performance. *Amer. Psychologist*, 1957, **12**, 458. (Abstract)

Hull, C. L. The rat's speed of locomotion gradient in the approach to food. *J. comp. Psychol.*, 1934, **17**, 393–422.

HULL, C. L. *Principles of behavior.* New York: Appleton-Century, 1943.

HULL, C. L. *A behavior system.* New Haven: Yale Univer. Press, 1952.

KNARR, F. A. Performance as a function of differential consummatory activity. Unpublished master's thesis, University of Missouri, 1959.

KNARR, F. A., & COLLIER, G. Taste and consummatory activity in amount of reinforcement effects. Paper read at Midwestern Psychological Association, St. Louis, April 1960.

PEREBOOM, A. C. A note on the Crespi effect. *Psychol. Rev.*, 1957, **64**, 263–265.

PEREBOOM, A. C. A goalless gradient. *J. exp. Psychol.*, 1958, **55**, 31–33.

SPENCE, K. W. *Behavior theory and conditioning.* New Haven: Yale Univer. Press, 1956.

WIST, E. R. Amount, delay, and position of delay of reinforcement as parameters of runway performance. Unpublished doctoral dissertation, University of Missouri, 1960.

(Received November 21, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 2, 167–176

THE INFLUENCE OF EARLY EXPERIENCE ON THE FRUSTRATION EFFECT [1]

ABRAM AMSEL AND ELIZABETH C. PENICK
University of Toronto *Washington University*

This experiment was designed to test the influence of early experience and training on the frustration effect (FE) as defined by Amsel and others (Amsel & Hancock, 1957; Amsel & Roussel, 1952; Wagner, 1959). It arises, primarily, out of a difference in interpretation of the phenomenon. While some workers in this area (e.g., Amsel, 1958; Brown & Farber, 1951) have leaned to a nonassociative interpretation of the FE, others (e.g., Lawson & Marx, 1958; Marx, 1956) have suggested that an associative mechanism accounts for it equally well. While the two points of view are not incompatible—associative and nonassociative factors might certainly operate together in the FE—it seems desirable to initiate some experiments to determine the extent to which variations in early experience factors, which can only affect later behavior associatively, determine the extent of the FE, as it has usually been measured in the double-runway apparatus. The investigation involved three experimental variables.

Rationalizing the first experimental variable, the number of pups to the litter in which *S* was reared, it is reasonable to suppose that in a large litter (12 pups) greater activity is required in securing food than is required in a small litter (3 pups); observation of large and small litters in the act of feeding demonstrates this. The differential amount of activity required of each type of litter in relation to temporarily unavailable food could be ascribed to several factors: the amount of space around the mother, the availability of a functioning nipple, and the amount of food available. It seems reasonable to suppose, therefore, that in the period before weaning large-litter animals are more likely to be frustrated when hungry, and are more likely to be active in response to frustration than animals raised in small litters, who are less likely to encounter factors inducing frustration in relation to feeding, and to learn vigorous activity in this connection.

The second experimental variable, conditioning of activity and inactivity to the frustrative presentation of food, was introduced after weaning, and was a further attempt to manipulate the early experience of *S*s. Some of the food-deprived *S*s were trained to vigorous activity in the presence of inaccessible food, while others were given food only after a period of relative inactivity. It appears reasonable to suppose, on the basis of a purely associative interpretation, that those *S*s trained to be inactive in the presence of inaccessible food would manifest a frustration effect, at later testing, of a lesser magnitude than those *S*s trained to be active in the presence of inaccessible food.

The third experimental variable was introduced during the test period in the double runway on nonreward trials. Half of the *S*s found an empty food cup in G_1 on nonreward trials, whereas the other half found food

[1] This experiment was performed at Tulane University and was supported by a grant, G-5527, from the National Science Foundation.

present but inaccessible (a wire covering allowed *S* to see and smell the food but not eat it). This variable may be rationalized as follows: food-not-present in G_1 is the usual frustration-defining situation; food-present-but-inaccessible in G_1 is a situation more similar to the frustrative feeding experiences of *S*s in the previous training stage, and it might be expected that there would be greater "transfer" of learned vigor or quiescence to this frustrative situation than to the food-absent situation.

If the frustration effect is a result of, or is strongly affected by, learning which has occurred prior to the actual testing situation, then the experiment should support the following generalizations: *S*s which learned to respond to nonreward situations by increases in activity, particularly those from large litters, should when tested in the double-runway apparatus, particularly under the food-present-but-inaccessible condition, show the greatest frustration effect. Under the same test conditions, *S*s which learned early to respond to nonreward situations with a decrease in activity, particularly those from small litters, should show the smallest frustration effect.

Method

Subjects

The *S*s were 40 hooded rats from the colony maintained at Newcomb College, Tulane University. Two *S*s died just before testing. The *S*s were tested when approximately 107 days of age, although the experimental procedures began within a few days after their birth.

Apparatus and Procedure

The entire experiment was conducted in four stages: (*a*) Birth of *S*s to weaning; (*b*) isolation and activity-inactivity treatments; (*c*) preliminary adjustment to double runway, and test in activity cages under stimulus change conditions; (*d*) test in the double-runway apparatus. Each stage will be described separately both with respect to apparatus and procedure employed.

Stage 1.—This stage extended from birth of *S*s to weaning at approximately 31 days of age. Ten mother rats produced 58 pups. When approximately 4 days old, they were taken from their biological mothers and a split-litter technique was used to separate the pups into two groups of 24 each. Two litters of 12 pups each were formed from one group and eight litters of 3 pups each were formed from the other. Between the two litter-size conditions control for original litter differences was exercised as far as was possible. Of the 24 rats raised in a large litter, 20 were randomly selected when approximately 31 days of age for the large litter (L) condition. Of the 24 rats which were raised in a small litter, 20 were randomly selected for the small litter (S) condition.

Stage 2.—Stage 2 extended from weaning at Day 31 to Day 102. When *S*s were approximately 31 days of age, they were separated and put into individual living cages. The housing construction, composed of $6\frac{1}{2} \times 11 \times 7$ in. cages, was in a separate room, dimly and indirectly illuminated at all times by two $7\frac{1}{2}$-w. bulbs. For 9 days *S*s were allowed to adapt to their home cages and were maintained on an ad libitum diet for food and water. The *S*s were then placed on $11\frac{1}{2}$ hr. food deprivation with water present at all times. Four days were allowed for *S*s to adjust to the new feeding schedule and become adapted to the frustrative-feeding apparatus.

The frustrative-feeding apparatus was a $1 \times 1 \times 1$ ft. box of unpainted plywood with a clear glass covering. Placed 1 ft. above the box was a 15-w. bulb which made the apparatus much brighter than *S*s' home cages. A wall separated *E* and the home cages from the apparatus, and during the frustrative-feeding treatments *E* observed *S*s' behavior through a 2-in. hole in this wall and a mirror arrangement. Another small opening in this wall permitted *E* to insert a forceps into the interior of the apparatus to manipulate the food in the feeding treatments.

During a 4-day period of adaptation to the feeding apparatus, each *S* was put into the box twice a day for 2 min., with the forceps inserted. No food was held in the forceps at this time. Upon completion of the 2-min. adaptation period, *S* was placed in its living cage where it found a 4-gm. pellet of Purina laboratory chow, and was allowed to eat as much of the pellet as was possible in $\frac{1}{2}$ hr.

Upon completion of the adaptation period, the frustrative-feeding procedures began.

Each of the two groups, L and S, was randomly divided into two subgroups, making four groups with 10 *S*s in each. One of the two subgroups from Group S and one from Group L were trained to be inactive in response to the frustrative-feeding situation (Cond. I); the other subgroups from Groups S and L were trained to be active (Cond. A).

In Cond. I, a single *S* was put into the feeding apparatus. A 1-gm. piece of chow, held in a pair of forceps, was inserted into the feeding box and *S* was allowed to pull at the food and attempt to bite into it initially, but was not allowed to have and eat the pellet until its gross overt activity had ceased. The period of required inactivity was gradually increased until *S* remained inactive in the presence of food for at least 6 sec. In Cond. A, *S* was treated in the same manner as above, except that it was allowed to obtain the food only while struggling vigorously for it.

After each trial, *S* was returned to its home cage and found food. For the first 12 days, while *S* was on an 11½-hr. deprivation schedule, frustrative-feeding procedures were carried out twice a day after which *S* received a 4-gm. pellet for ½ hr. After this, when it was thought that *S*s could survive longer periods of deprivation, a 23-hr. feeding schedule was instigated and maintained throughout the remainder of the experiment. The frustrative-feeding procedures were then administered once a day, after which *S* found two 4-gm. pellets, and was allowed to eat for a period of 1 hr.

Stage 3.—In order to evaluate the effectiveness of procedures of Stage 2, all *S*s were run through an independent test of activity on Days 102 to 107. In particular, interest was focused on the possibility that "spontaneous activity" could be affected by a generalized form of the stimulus change (increase of illumination) associated with active or inactive behavior established in the prior stage of training. However, the major stimulus to activity or inactivity, temporarily inaccessible food, was not involved in this test.

Groups of 10 *S*s were brought in carrying cages into a windowless interior room which contained 10 individual stabilimeter-type activity cages of a design adapted from Campbell and Sheffield (1953). This apparatus has recently been described in detail by Amsel and Work (1961). Each cage sat in what was essentially an open-type box of white Celotex in order to provide a nearly homogeneous visual field and to minimize intercage visual and auditory stimuli. The range of temperature in the course of this portion of the experiment was 78° to 81°. The cages were arranged in a circular array below two ceiling lights. The first light, illuminated during a 1-hr. adaptation period was 7½ w.; the second light which produced the sudden stimulus change was a 200-w., semi-silvered bulb.

The *S*s were placed individually in the 10 activity cages for a 1-hr. adaptation period under the dim (7½ w.) illumination. After this period, activity was recorded for 5 min. under dim illumination in order to obtain the prestimulus-change, basal activity levels. The bright (200 w.) overhead light was then switched on, and activity was recorded in 1-min. intervals for 5 min. The 1-min. intervals were separated by 15-sec. intervals during which time the counters were not in operation, which afforded an opportunity for *E* to record the activity for the previous minute.

At the termination of the activity test each day, *S*s were removed from the activity cages, placed in the carrying cages, and brought into a room containing the double-runway apparatus (described in detail later). Preliminary adjustment to this apparatus was carried on during this 6-day period. Each *S* was allowed to explore the double runway singly for 5 min. with all doors down, and the photoelectric system operating. They were then returned to their home cages.

All 40 *S*s were tested in this manner. In view of the fact that only 10 *S*s could be tested at one time, a latin square design was used in order to control for any possible effects the time of testing in the activity cages may have had. Since this test was made in the first 4 successive days of this 6-day period each squad of 10 *S*s was tested in all of the four necessary periods which extended from 6 hr. to 1 hr. before the regular feeding treatments.

Stage 4.—When *S*s were approximately 107 days old, the frustrative-feeding treatments were discontinued and testing for magnitude of the FE was begun. Each of the four groups was randomly divided into two further subgroups. One of these groups found the food cup empty on nonreward trials; the other found the food cup covered with a screen covering which allowed *S*s to see and smell food in the cup but not to eat it. This variable was labeled FN (food not present) and FP (food present but inaccessible).

The double-runway apparatus consisted of an L shaped double runway (See Fig. 1 for

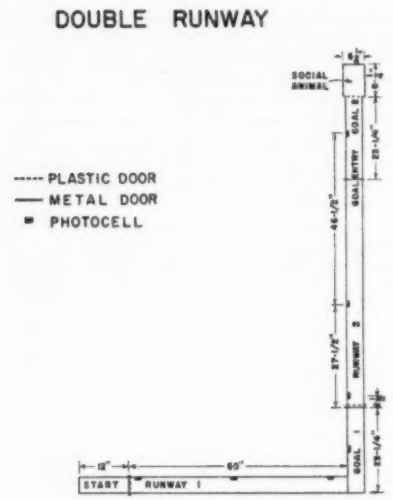

FIG. 1. Floor plan and inside dimensions of double runway.

dimensions). The first section was composed of a start box (SB), Runway 1 (R_1), and Goal Box 1 (G_1); the second section began with G_1 which served as the start box for the second runway, Runway 2 (R_2), and Goal Box 2 (G_2). It was necessary to divide G_2 into two compartments, one of which served as a temporary waiting box for a "social" *S* while the other served as a common restraining box for the social and experimental *S* when *S* entered G_2. A small, black, $1 \times 1\frac{1}{2}$ in. food cup, $1\frac{1}{2}$ in. from the floor and approximately 2 in. from the start door of G_1, was present on all trials. A screen cover was made to slip on and off the food cup. The walls of the SB, R_1, and G_1 were black and the flooring of these sections was black rubber mat. The walls of R_2 were black and white alternating vertical stripes, and the floor was $\frac{1}{4}$-in. brass bars spaced $\frac{1}{2}$ in. apart. Box G_2 was white. The top of the apparatus was $\frac{3}{16}$-in. Plexiglas, hinged to allow access to the runways. The Plexiglas over the first section of the apparatus was overlaid with red celluloid sheeting to decrease further the level of illumination in the R_1 relative to R_2. Illumination was provided by two, 40-w. fluorescent lamps suspended in series approximately 22 in. above R_2.

The procedure followed for the acquisition of the running responses under continuous reward was essentially that of Amsel and Roussel (1952). For 18 days, each *S* was given three trials per day in the apparatus under 23 hr. hunger. During the continuous reward trials food was present at all times in G_1, and a "social" *S* was present in G_2. The details of procedure on an individual trial are as follows: (*a*) *S* was put into the SB and, after it was oriented to the door, a metal door was dropped; (*b*) 3 sec. after the door was dropped, a transparent plastic door was dropped; (*c*) *S* traversed R_1, entered G_1, and ate two small pellets (about 1 gm.); (*d*) after a 20-sec. interval, controlled automatically by an interval timer, a metal orienting door was dropped followed again by the dropping of a Plexiglas door after *S* oriented in that direction; (*e*) *S* traversed R_2 into G_2 and was allowed contact with the social *S* for 20 sec., after which *S* was removed and returned to the carrying cage. The intertrial interval was not less than 20 min.

Performance measures were taken in the second section of the apparatus by means of Standard Electric timers activated by a photoelectric system. The sequence of starting and stopping the clocks was as follows: (*a*) Twenty seconds after a photoelectric unit in the center of G_1 had been activated, a solenoid arrangement activated by an interval timer caused a metal door to drop. This was followed by the dropping of the Plexiglas start door in G_1 which started Clock 1. (*b*) When *S* moved to a point $2\frac{1}{2}$ in. into R_2 Clock 1 stopped, providing a measure of latency (L), and Clock 2 started. (*c*) Passage 25 in. down R_2 stopped Clock 2, providing a measure of running time (RT), and started Clock 3. (*d*) Entrance midway into the first section of G_2, $46\frac{1}{2}$ in. away, stopped Clock 3, providing a measure of goal-entry time (GT).

The 18 days of continuous reward training were followed by the test period during which each *S* continued to run three trials a day for 12 days. Of these, half were nonreward (N) trials and half were reward (R) trials. The nonreward trials differed from the reward trials in that *S* found either an empty food cup or a covered food cup containing two small pellets in G_1. Contact with the social *S* in G_2 for 20 sec. continued to be the reward for traversal of R_2. The following arrangement of daily trials was employed to control for any possible effects due to the order: NRN RNR NNR RRN NRR RNN NRN RNR NNR RRN NRR RNN. Over each 2-day block of six trials, there were three N and three R trials, one of each kind in each trial position.

Results

Test of Activity to the Generalized Activity Cue

All four groups showed activity increases to the change in illumination on each test day except for a single point, the third day for Group S-I (Fig. 2). The I groups, for both small and large litters, were generally less active than the A groups. Increases in activity seem greater in the L groups than in the S groups. The prestimulus-change levels of activity are about the same for both the S vs. L and the A vs. I variables. Averaging over all 4 days, the stimulus-change activity levels were: S-A, 10.9; S-L, 8.0; L-A, 12.7; and L-I, 9.7. The prestimulus-change values in the same order were 6.3, 4.6, 5.3, and 4.5. In an analysis of variance of the prestimulus-change measures, none of the main effects was significant, F for between groups being less than unity. In the stimulus-change analysis the between-groups F was significant beyond the .01 level, but neither of the main effects (A vs. I, L vs. S) was significant at the .05 level. The interaction F was 4.68, significant beyond the .01 level. An analysis of covariance testing for group activity differences during the stimulus-change period adjusting for prestimulus-change differences produced an identical pattern of significance relationships. In view of the extremely large variability obtained between *S*s and the relatively small number of *S*s in each group, one might hesitantly conclude that treatments established in Stage 2 of this experiment had some effect upon the reaction to a stimulus change (illumination change) in Stage 3.

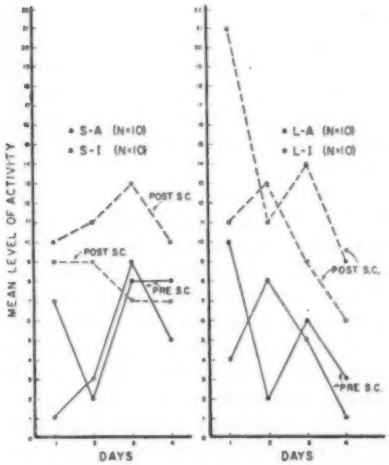

FIG. 2. Activity in stabilimeter test situation before and after stimulus change.

Test of the FE in the Double Runway

Preliminary to an analysis of variance to test differences in magnitude of the FE, t tests for related measures were performed to demonstrate that there was a significant FE in each of the eight groups. A median running time score over 2-day blocks of trials for R and for N trials was found for each *S* over the test period. The means of the median running times obtained on the last 6 days of testing for reward and nonreward trials were used in this analysis. In every case the difference was in the direction of the FE, and values of t ranged from 2.39 to 4.70, significant in each case beyond the .05 level in two-tailed tests.

Relation of FE Magnitude to Experimental Variables

Although this study was originally designed as a $2 \times 2 \times 2$ factorial, several considerations prevented a very meaningful comparison of the magnitude of the FE in small- and large-litter groups. First, the result of an analysis of variance of the weights of *S*s 1 week before testing

showed the small-litter *S*s to be significantly heavier than the large-litter *S*s. There was no significant difference in weight attributable to the frustrative-feeding (A vs. I) variable, nor was the interaction between litter size and the A vs. I feeding procedures significant. Secondly, during the introduction of the feeding procedures in Stage 2 the large-litter *S*s of Cond. I required a longer period of time to reach the 6-sec. inactivity criterion than did the small-litter *S*s; and the large-litter *S*s were more readily trained to vigorous struggling than were the small-litter *S*s upon presentation of food in Cond. A. Thirdly, as Fig. 3 and 4 show, R_2 running time during the testing was very much slower on both R and N trials for small-litter *S*s than for large-litter *S*s, although the FE shows up clearly for both groups, a finding which holds true for latency data and for goal-entry data as well. In view of the above, and the attendant possibility that the litter-size variable is confounded with a weight variable, separate statistical analyses were made for the large- and small-litter groups.

Two analyses of variance, one for L and one for S groups, were performed using the mean median running time on N trials in R_2, which for each *S* was a mean of six, 2-day (3-trial) medians. The justification for using N-trial scores alone as indicants of FE for intergroup comparisons was a between-groups analysis of R scores which yielded nonsignificant between-groups *F* ratios for R trials for both the L and S conditions. In order to equate the number of *S*s per cell of the analysis, an arbitrary *S* was added to each of the two groups in which an *S* died by using a table of random numbers and selecting the scores of 1 experimental *S* over the entire test period. A nonsignificant *F* ratio was found in both analyses for between-group means. However, the Trials × Groups interaction was significant beyond the .05 level in both analyses, and the between-trials *F* ratio was significant for the L groups, but not for the S groups.

The implications of the above analyses can be clarified by a graphic analysis of the data. Figure 3 shows a breakdown for the large-litter groups of mean median running time on R and N trials over the 12 days of testing. The two top panels compare Cond. A and I for all large-litter *S*s, holding constant the variable introduced in Stage 4 of the experiment (FN and FP treatments). The two bottom panels compare the FN and FP treatments for all large-litter *S*s. While the levels of performance on N and R trials are highly similar for Cond A and I, suggesting that the procedures instigated in Stage 2 did not differentially effect the magnitude of the FE, these curves do show a reversal of the FE at the first points of Cond. A and I. This early reversal can be clarified by reference to Cond. FN and FP in the lower panels of Fig. 3, where the A and I procedures introduced in Stage 2 of the experiment are held constant. It would appear that magnitude of the FE is affected by the Stage 4 (FN vs. FP) procedure at the first test point: food present but inaccessible at first produces an increase in running time in R_2 following nonreward; food not present, which is our usual nonreward procedure, produces an immediate FE. (This effect shows up even more strikingly in the latency data, which are not presented.)

Figure 4 presents, for the S groups, the same graphic breakdown as has just been given for the L groups. Food present but inaccessible on

nonreward trials again produces an apparent reversal of the FE for the first few days, while the FE shows immediately for the group which found no food on nonreward trials. This once more accounts for the first point reversal of the FE in Cond. A, the virtual absence of FE at the first point in Cond. I., and for the significant Trials × Groups interaction.

As indicated earlier, an analysis of variance of the reward running times

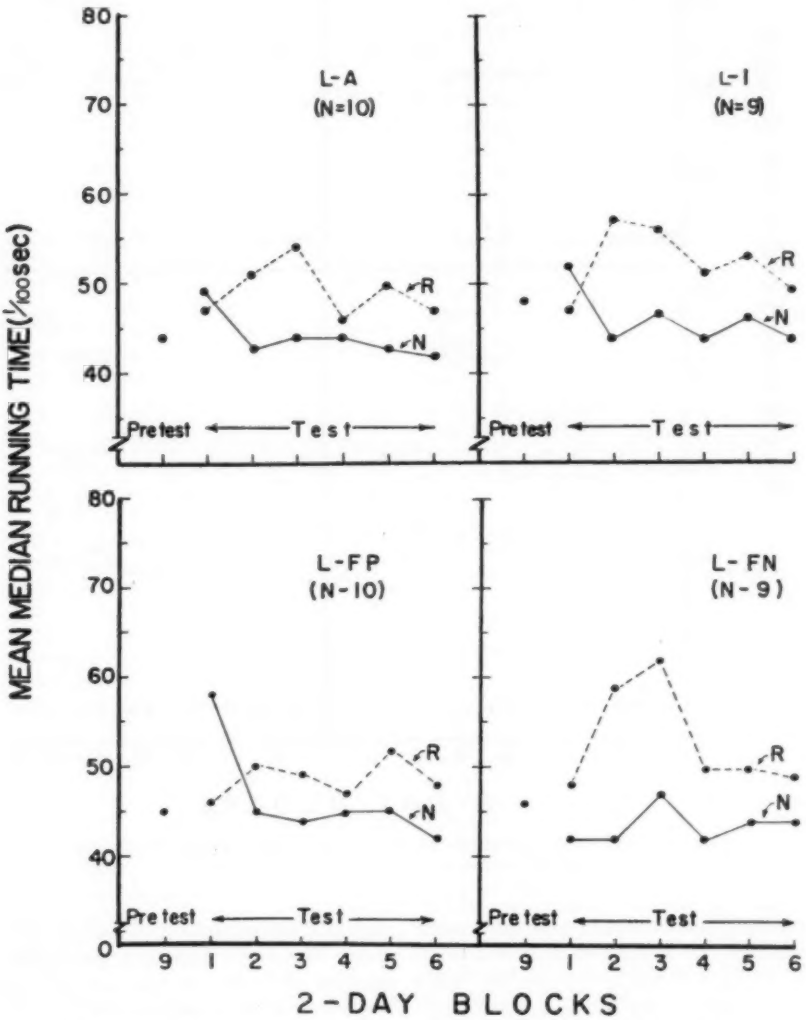

FIG. 3. Breakdown of the effects of the Activity-Inactivity variable and the FN-FP variable for the large-litter condition.

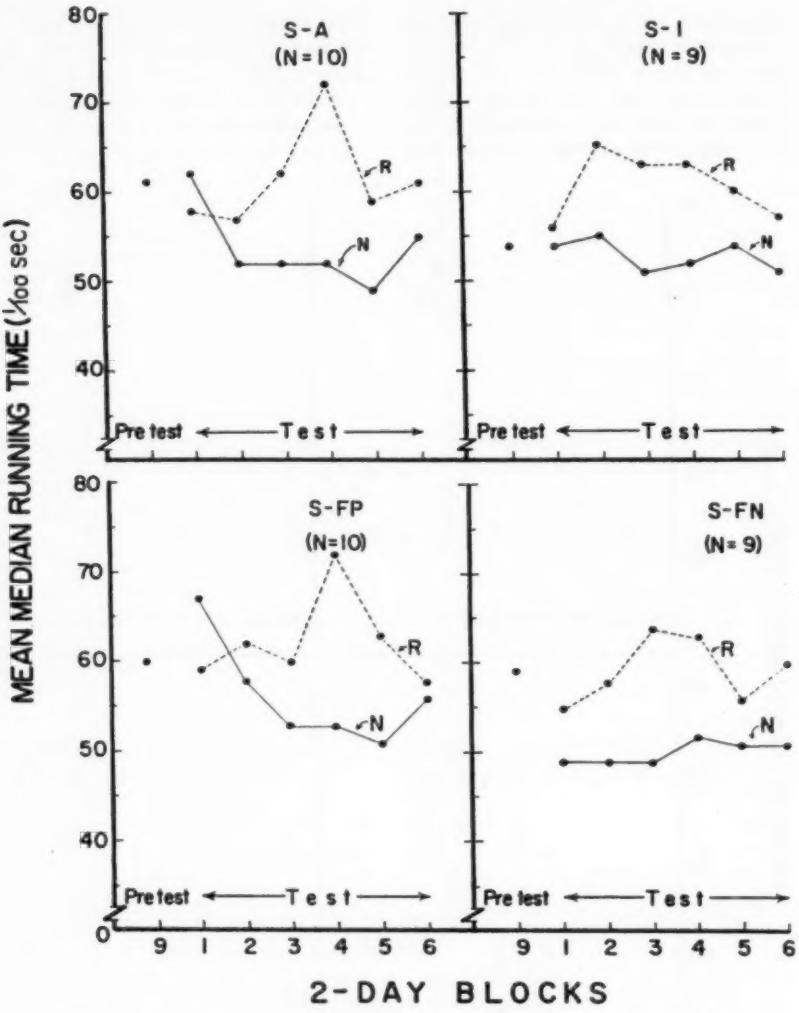

FIG. 4. Breadkown of the effects of the Activity-Inactivity variable and the FN-FP variable for the small-litter condition.

for both S and L groups over the 12-day test period yielded nonsignificant F ratios for between-groups and for Trials × Groups interactions. However, in both analyses, F was significant at the .95 level for the between-trials variation. It should be noted that in some cases the reward curves do appear to rise early in the test period and then fall again, a finding which is not characteristic of other data of this sort. A significant F ratio for be-

tween trials was also found in the analysis of nonreward scores for the L groups, and approached significance in the S groups. However, in the nonreward curves, the between-trial variation is more readily attributable to the decrements in running time on the first few trials in the FP groups.

Discussion

If the suggestions made by Marx (1956) as to the associative nature of the FE have been understood correctly, it would appear that, at the very least, such an interpretation would predict differences in the magnitude of the FE between the A and I groups, and that it might even predict no FE or even a reversed FE in those groups which were trained to remain still in the presence of inaccessible food when hungry, particularly in the case of *S*s reared in litters of three. It may be argued, of course, that the procedures defining the variables introduced in this experiment were not adequate to demonstrate the associative determination of the FE. However, Marx suggests no specific procedures to follow, and the method followed here was to provide as divergent associative experience as possible in relation to food-getting beginning at birth. A generalized drive interpretation (e.g., Amsel, 1958) of the FE still seems tenable, although this experiment is no specific test of such a position.

The apparent reversal of the frustration effect found in the FP groups on the first few trials is readily understood in terms of competing responses. The *S*s of this condition, upon finding the food cup covered on the first few nonreward trials, vigorously attempted to bite and claw through the screen covering. This was true of the inactivity-trained as well as the activity-trained *S*s, and was often accompanied by vocalization, urination, and defection. Some *S*s remained in G_1 for as long as 4 min. after the dropping of the metal door before orienting toward R_2; and on several occasions *S*s would pass the first photocell starting the running time measure and then return to G_1. This behavior is clearly reflected in the latency data for FP groups (not presented) which show very marked increases in the group means and variability on the first few nonreward trials. As these responses disappeared, the increased running speed in R_2 after nonreward trials was clearly evident. The *S*s of Cond. FN, on the other hand, showed none of the above behavior, and the latency data of these groups are much more stable.

The implication of these observations relative to the FP-FN variable is obvious. While we have defined the FE as a measurable increase in response vigor in the second runway of the double-runway situation, and this has been our major dependent variable, there are certainly other possible effects of frustrative nonreward. The FP condition points up the fact that there can be frustrative effects which appear to reduce the criterion FE or prevent it from manifesting itself, at least temporarily. This does not mean that there is not increased *D*, but only that there are for a time stronger *H*s which this increased *D* is energizing. This produces the appearance of decreased strength of the criterion (FE) response. As soon as these competing tendencies (vigorous behavior in relation to the inaccessible food) weaken, the FE which has been "masked" becomes apparent.

Summary

The major concern of this investigation was to test the effect of early experience in relation to food-getting on the magnitude of the frustration effect. Three independent variables were manipulated from the birth of 40 hooded rats until the time of testing in a double-runway apparatus. These variables were: the size of the litter in which *S* was raised; differential frustrative-feeding procedures carried out in a feeding apparatus clearly discriminable from *S*'s home environment; and no-food-present vs. food-present-but-inaccessible in the first goal box of the double-runway apparatus on nonreward trials.

A preliminary test, involving a stabilimeter apparatus, was run to evaluate the effective-

ness of the activity-inactivity training procedure.

The results of the main test in the double runway indicated the following: (*a*) Training hungry *Ss* to be active or inactive in relation to food-getting does not appear to "transfer" to the double-runway situation in such a way as to influence the magnitude of the FE either in *Ss* raised in large or in small litters. (*b*) During the testing period, food-not-present on nonreward trials produced an immediate and continued FE in all groups run under this condition (*c*) The presence of inaccessible food on nonreward trials appears to produce an initial increase in running times, and an apparent reversal of the FE in Runway 2. This reversal is temporary and by Day 6 of the testing period the usual FE is clearly present in all groups of this condition.

References

Amsel, A. The role of frustrative nonreward in noncontinuous reward situations. *Psychol. Bull.*, 1958, **53**, 102–119.

Amsel, A., & Hancock, W. Motivational properties of frustration: III. Relation of frustration effect to antedating goal factors. *J. exp. Psychol.*, 1957, **53**, 126–131.

Amsel, A., & Roussel, J. S. Motivational properties of frustration: I. Effect on a running response of the addition of frustration to the motivational complex. *J. exp. Psychol.*, 1952, **43**, 363–368.

Amsel, A., & Work, M. S. The role of learned factors in "spontaneous" activity. *J. comp. physiol. Psychol.*, 1961, **54**, 527–532.

Brown, J. S., & Farber, I. E. Emotions conceptualized as intervening variables—with suggestions toward a theory of frustration. *Psychol. Bull.*, 1951., **48**, 465–495.

Campbell, B. A., & Sheffield, F. D. Relation of random activity to food deprivation. *J. comp. physiol. Psychol.*, 1953, **46**, 320–322.

Lawson, R., & Marx, M. H. Frustration: Theory and experiment. *Genet. Psychol. Monogr.*, 1958, **57**, 393–428.

Marx, M. H. Some relations between frustration and drive. In M. R. Jones (Ed.), *Nebraska symposium on motivation: 1956*. Lincoln: Univer. Nebraska Press, 1956.

Wagner, A. R. The role of reinforcement and nonreinforcement in an "apparent frustration effect." *J. exp. Psychol.*, 1959, **57**, 130–136.

(Received November 22, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 2, 177–182

THE PROGRESSION-REGRESSION HYPOTHESES IN PERCEPTUAL-MOTOR SKILL LEARNING[1]

ALFRED H. FUCHS[2]

Ohio State University

In a recent experiment Garvey and Mitnick (1957) demonstrated that perceptual-motor skill development may be described analytically by treating human information processing as analogous to that of a servosystem. Their analysis of lead-lag tracking errors suggested that the human operator in a continuous closed-loop control system behaves progressively, with training, as a more complicated servomechanism. The course of this progression appeared to be systematic: initially, responses to displayed system error were based on the *velocity* aspects of that input signal (i.e., its rate of change), while later in training the more subtle error acceleration and even Δ acceleration (jerk) aspects determined the control responses. This systematic development has been termed the *progression hypothesis* by Fitts, Bahrick, Noble, and Briggs (1961). This hypothesis holds that in perceptual-motor skill development *S* learns to respond to higher derivatives of the error signal's amplitude (e.g., velocity and acceleration characteristics) and to *weight* these selectively in achieving system stability and error minimization. This extension of the Garvey and Mitnick (1957) results follows from derivations of transfer functions for the human operator. Such transfer functions are descriptive of the transformations and translations imposed by the human operator on input information (system error) in generating his output (force applied to a control device). Such research, deriving human transfer functions through spectral and/or correlational analysis of tracking records, has indicated that error amplitude is more heavily weighted than error velocity in a simple positional control system (Russell, 1951, Walston & Warren, 1953), while in a more complex control system (velocity or first-order control dynamics) error velocity receives greater weight (Tustin, 1947). This implies that the initial level of control exercised by the human is the minimum necessary level to achieve system stability and that this minimum level is determined by the complexity of the system.

Further evidence for the validity of the progression hypothesis derives from research, carried out at the Goodyear Laboratories (Goodyear Aircraft Corporation, 1955), which demonstrated that two pilots of differing experiential background weighted velocity and acceleration characteristics of an error signal more heavily than error amplitude; however, only the more highly experienced *S* weighted the acceleration characteristic more

[1] This report is based on a dissertation submitted to the Graduate School of the Ohio State University in partial fulfillment of the requirements for the PhD degree. The author is indebted to George E. Briggs for his guidance throughout the course of the study, which was supported in part by Contract No. AF 30(602)-2107 between the United States Air Force, Rome Air Development Center, and the Antenna Laboratory, Department of Electrical Engineering of the Ohio State University. Permission is granted for reproduction, translation, publication, use, and disposal in whole and in part by or for the United States Government.

[2] Now at Electric Boat Division, General Dynamics Corporation, Groton, Connecticut.

heavily than error amplitude. Perhaps the most significant contribution of the Goodyear research was the direct technique employed to "fit" the human transfer function: a differential equation plus equation expressions for "neuromuscular" lag and for such nonlinearities as human thresholds and "dither" were placed in an analog computer. The *E* manipulated the human analog constants until the analog output matched the human output. The results of this matching were remarkable in that when the analog was substituted for the man in the control system, *S* (who continued to "control" unaware that he had been replaced) could not detect the substitution for periods of 20 to 30 sec. of tracking!

A corollary of the progression hypothesis is the regression hypothesis (Fitts et al., 1961) which predicts that if following skill development *S* experiences task-induced stress, weight assigned to the higher derivatives of displayed error will decrease and weight assigned to error amplitude will increase, i.e., *S* will "regress" or return to a less complex level in controlling the system. This hypothesis was suggested by the data of Russell (1951) who showed that as complexity of the input increased, relative weight assigned to error amplitude also increased. This hypothesis has received attention by Miquelon (1958) who utilized the Garvey and Mitnick (1957) technique of analyzing lead-lag tracking error scores as a means of testing the regression hypothesis. While his results were interpreted as supporting the hypothesis, they were not unequivocal. An adequate test of both the progression and regression hypotheses requires the measurement of the weights assigned to the stimulus (displayed error) characteristics and the empirical derivation of their changes during periods of training and of task-induced stress. This paper is a report of such a test.

Method

Apparatus.—The experimental technique was similar to that at the Goodyear Laboratories (Goodyear Aircraft Corporation, 1955) in that an analog was utilized to approximate the information-processing characteristics of the human *S* in a continuous control task. The analog was constructed to operate according to a simplified version of the human transfer function utilized in the Goodyear research (McRuer & Krendel, 1957).

The analog consisted of a differential equation of the form:

$$\theta_e + \dot{\theta}_e + \ddot{\theta}_e = (K_1\xi + K_2\dot{\xi} + K_3\ddot{\xi})e^{-\tau p}$$

where θ_e is the human output, ξ is displayed tracking error, $e^{-\tau p}$ is an exponential lag corresponding to human reaction time in continuous control tasks, and K_1, K_2, and K_3 are weighting constants for positional, velocity, and acceleration aspects of the error signal. This equation was set up on a Pace analog computer (Model 16-31R) and the servounits of a Reeves analog computer (Model C-10) hunted optimum settings for K_1, K_2, and K_3 during the tracking trials. Thus, the technique may be described as an on-line fitting of the above equation to human performance. Compared to off-line techniques previously used, the Goodyear approach is a remarkably efficient means of deriving transfer functions. A flow diagram of the system is provided in Fig. 1. The signal E^2 provided an input to the servounits and thus a servo continued to hunt for an optimum K_1, K_2, or K_3 until E^2 was reduced to zero. Since E^2 always existed at some finite level, the servo never stopped but would oscillate around the optimum value which was recorded as the value of a weighting constant for a given trial. During a particular trial only one weighting constant was estimated, the values of the other two being set at the

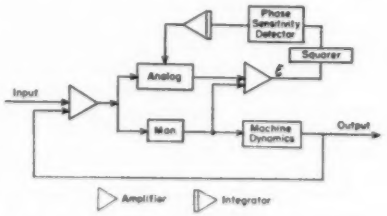

FIG. 1. The experimental system.

average of their determined values from previous trials.

Subjects and procedure.—Five undergraduate male volunteers served as *S*s and all were paid for service in the experiment. The *S*'s primary task was that of compensatory tracking; the display was an oscilloscope, the control was a spring-centered control column, and the machine dynamics through which *S*s tracked consisted of two analog integrations. A random signal provided the input, and the difference between this signal and *S*'s control movements comprised the error signal shown on the display. The *S*s were instructed to keep the cursor on target; their ultimate goal was to achieve and hold zero error.

Throughout the training period (20 days) *S* received nine trials per day; each trial had a duration of 120 sec., with an intertrial interval of 60 sec. Knowledge of their performance was accessible to all *S*s throughout the 25-day duration of the experiment; it was hoped that this would facilitate and maintain their motivational level.

After the training period, the final 5 days of the experiment constituted a transfer condition in which *S* was required to perform a secondary task simultaneously with the above-mentioned primary task; this condition is defined as 'task-induced stress" (Lazarus, Deese, & Osler, 1952). The secondary task was a simple positional tracking system (Gain & Fitts, 1959) in which *S* attempted to track a .06-cpm input on a compensatory display. Instructions to *S* at the start of the transfer trials emphasized the necessity of doing as well as possible on *both* tasks. The secondary task was controlled by *S* with his left hand while tracking in the primary task with his right hand.

Performance measures utilized were (*a*) system error produced in the primary task by *S* (difference between the input and the system output), (*b*) *S*'s error on the second task, and (*c*) the gain settings used by the analog for the position, velocity, and acceleration constants of the transfer function which enabled it to match the output of *S*. The *S*'s error was not scored for the first 10 sec. of the 120-sec. trial in order to eliminate initial transients from total trial error and thus achieve a more reliable measure of *S*'s performance. The first 5 sec. of performance error on the secondary task were not scored for the same reason. Error scores during training were summarized in 2-day blocks (18 trials); the gain settings were treated similarly (6 trials per gain setting). The transfer data were summarized over single days; 9 trials for error scores, and 3 trials for each of the gain settings per day.

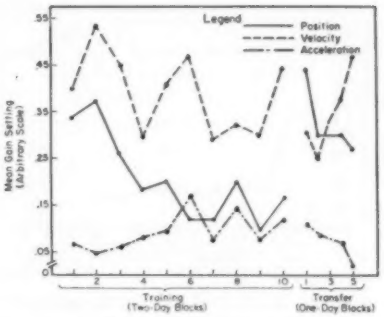

FIG. 2. Mean weights (analog gain settings) assigned to stimulus components during training and transfer.

RESULTS

Training.—The weights assigned to the stimulus components throughout the experiment are summarized in Fig. 2. The trends during training are evident: position gain, representing the weight assigned by *S*s to error amplitude, decreased during training while acceleration gain increased; the weight assigned to the velocity component of the stimulus fluctuated around a line having a generally downward trend. Analysis of these data (Table 1) reveals that these scores not only differed with respect to one another, but they also changed differentially as a function of training, as seen in the significant ($P < .01$) Measures and Measures × Days effects. The significant ($P < .01$) Sub-

TABLE 1

ANALYSIS OF VARIANCE OF THE TRAINING DATA

Source	*df*	*MS*	*F*
Measures (M)	2	11,112.495	104.58**
Blocks	9	368.332	2.14*
Ss	4	188.650	2.41
M × *Ss*	8	106.255	1.36
M × Days	18	241.782	3.08**
Days × *Ss*	36	171.983	2.19**
Days × *Ss* × M	72	78.398	

* $.05 < P < .10$.
** $P < .01$.

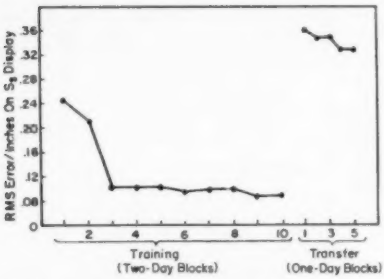

FIG. 3. Amount of error on the primary tracking task.

jects × Days interaction indicates, of course, different rates of change for these *Ss* on each of the measures employed.

The usual indicator of tracking performance, the RMS error, is summarized in Fig. 3; it is readily apparent that *Ss* increased their proficiency of performance on the primary tracking task. The correlation (multiple *r*) between the weights given the stimulus characteristics and RMS error was .864 ($P < .05$). The only significant single correlation was between position gain and RMS error; it was found to be .858 ($P < .01$), an interesting finding since it corroborates the prediction of the progression hypothesis that perceptual-motor skill development consists of the reduction of the weight assigned to the position component of the stimulus, with a resultant *relative* increase in the weights assigned to the velocity and acceleration characteristics.

The weights (gains) accorded velocity and acceleration characteristics of the input were expressed as ratios of position gain in order to assess relative change during training; the asymptote of these plotted changes revealed a position/velocity ratio of .34 and a position/acceleration ratio of 1.02.

Transfer.—The shift in assigned weights from the training period to the transfer condition was analyzed (Table 2) and found to be significant ($P < .01$ for the Measures × Training-Transfer interaction). Despite the obvious relative shift in velocity and position weights *during* transfer, the Measures × Days interaction failed to attain statistical significance (Table 3); however, the over-all difference among measures was significant ($P < .01$). That *Ss* performed, and did not

TABLE 2

ANALYSIS OF VARIANCE OF THE CHANGES IN STIMULUS COMPONENT WEIGHTS FROM TRAINING TO TRANSFER

Source	*df*	*MS*	*F*
Measures (M)	2	1443.97	11.16**
Ss	4	440.84	3.34
Training-Transfer (TT)	1	17.14	.28
M × *Ss*	8	129.34	2.13
M × TT	2	759.57	12.51**
Ss × TT	4	131.81	2.17
M × *Ss* × TT	8	60.73	

** $P < .01$.

TABLE 3

ANALYSIS OF VARIANCE OF THE CHANGES IN STIMULUS COMPONENT WEIGHTS DURING TRANSFER

Source	*df*	*MS*	*F*
Measures	2	5978.42	21.63**
Ss	4	392.04	2.31
Days (D)	4	103.18	.90
M × *Ss*	8	204.46	1.20
M × D	8	276.34	1.63
Ss × D	16	114.06	.67
Ss × D × M	32	169.91	

** $P < .01$.

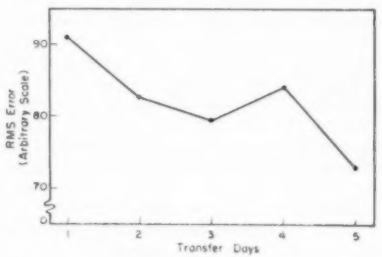

FIG. 4. Amount of error on the secondary tracking task.

ignore, the secondary tracking task is seen in Fig. 4; *Ss* improved on both tasks during transfer, and thus the transfer condition may indeed be considered to have involved task-induced stress.

Discussion

The training results conform to the predictions made from the progression hypothesis; at the outset of training the amplitude and velocity components of displayed error are given almost equal weight in the determination of the control responses. The decrease in position weighting, relative to velocity, indicates an early stage in the progression to a higher order of control, one in which the velocity characteristics of displayed error play an increasingly greater role in determining the response relative to the amplitude of displayed error. The lowest level of control, one in which error amplitude is the largest single determinant of the response, is not evident in these results, since the adoption of such a basis for control responses would undoubtedly have led to instability in the system. The validity of this assertion may be seen in the transfer data; on those days wherein error amplitude (position) is weighted relatively and absolutely greater than velocity, the RMS error score is higher than at any other time. The size of this error is even somewhat greater than Fig. 3 indicates, since on many of the initial transfer trials, system error surpassed the limits of the scoring mechanism; on those trials the maximum possible score was arbitrarily assigned, thus artificially limiting the size of error indicated in Fig. 3.

The relative increase in weight accorded error acceleration is even more dramatic, in that early in training it is negligible, while at the conclusion of training it approximates the weight assigned to error amplitude. That it failed to become *greater* than error amplitude is probably explicable on the basis that equivalence of these represents an optimal ratio for stability of this particular system.

Although direct comparability with previously determined values of the relative weights for error velocity and acceleration is somewhat limited since the means of deriving these weights and the systems in which they were utilized vary considerably, it is interesting to note that the ratios of the weights reported in the Goodyear (Goodyear Aircraft Corporation, 1955) study show some similarity with the values found here. Examination of Table 4 indicates that *Ss* of the present study (PR) fall between the values for the two pilots (VO represents an *S* with experience in flying light aircraft, while JB represents an *S* with considerable jet flight experience), indicating a level of control somewhat beyond the velocity order of pilot VO and approaching the acceleration level of pilot JB.

TABLE 4

Comparison of Ratios Reported by the Goodyear Laboratories and the Present Research

Pilot	Pos/Vel	Pos/Acc
VO	.71	7.14
JB	.03	.50
PR	.34	1.02

The transfer data reported in Fig. 2 substantiate the regression hypothesis; error amplitude receives greater weight during the initial days of transfer, after which the response is more largely determined by the velocity characteristic of displayed error. The decrease in the weight accorded the acceleration characteristic is probably to be accounted for by the stress condition itself. The small amount of available time to concentrate on the primary task prevented the utilization of this most subtle aspect of the display; the excessively high RMS error scores during transfer (Fig. 3) would seem to attest to this interpretation.

The direct support offered for these hypotheses also constituted indirect support for the assertion of Birmingham and Taylor (1954), Garvey and Mitnick (1957), and Fitts et al. (1961), that analogies between human performance and the performance of servosystems

can provide testable hypotheses in the growth of theory for perceptual-motor skill development. The conception of such development in terms of electronic functioning seems, in the light of the results of this study, a clearly valuable contribution in that hypothetical internal characteristics of the human controller are specified and open to empirical test. The specific technique necessary to test such hypotheses has shown itself to be not only feasible but remarkably efficient; it offers rapid approximations of subtle characteristics of human responding more sensitive than the usual measures of tracking performance. That such scores *are* more sensitive is apparent in the comparison of Fig. 2 and 3; weights accorded stimulus components continue to change beyond the point at which RMS error becomes asymptotic.

The conclusions to be drawn from this study seem apparent: perceptual-motor skill performance rests upon the development with practice of the performer's ability to perceive and weight more subtle stimulus cues in the determination of more adequate error-minimizing responses. The technique utilized in this investigation holds out the promise that an adequate theory of perceptual-motor skill learning may be built upon the empirical generalizations tested herein.

Summary

Two hypotheses regarding the development of perceptual-motor skill were tested. The first of these, the progression hypothesis, predicts that learning a perceptual-motor skill entails weighting more subtle aspects (velocity and acceleration characteristics) of the displayed error signal more heavily; the second, the regression hypothesis, predicts that under conditions of stress the relative weighting will "regress" to a heavier emphasis on more apparent stimulus characteristics (error amplitude), but that recovery from stress will evidence the properties predicted by the progression hypothesis. The technique employed to test these hypotheses was a system in which a human analog, consisting of analog computers and servomechanisms, was utilized to attempt to match the performance of S; the gain adjustments for the position, velocity, and acceleration characteristics (corresponding to the weights assigned by S in the system) were provided automatically by a servomechanism. These gain settings were treated as approximations of the weights assigned by S.

The results confirmed the hypotheses; the technique proved to be efficient and rapid, and it was concluded that the beginning of a theory of perceptual-motor skill learning may be found in these results.

References

Birmingham, H. P., & Taylor, F. V. A human engineering approach to the design of man-operated continuous control systems. *USN Res. Lab. Rep.*, 1954, No. 4333.

Fitts, P. M., Bahrick, H. P., Noble, M. E., & Briggs, G. E. *Skilled performance.* New York: Wiley, 1961.

Gain, P., & Fitts, P. M. A simplified electronic tracking apparatus (SETA). *USAF WADC tech. Rep.*, 1959, No. 59–44.

Garvey, W. D., & Mitnick, L. L. An analysis of tracking behavior in terms of lead-lag errors. *J. exp. Psychol.*, 1957, **53**, 372–378.

Goodyear Aircraft Corporation. Investigation of control "feel" effects on the dynamics of a piloted aircraft system. *Goodyear Aircraft Corp. Rep.*, 1955, No. GER 6726.

Lazarus, R., Deese, J., & Osler, S. The effects of psychological stress upon performance. *Psychol. Bull.*, 1952, **49**, 293–317.

McRuer, D. T., & Krendel, E. S. Dynamic response of human operators. *USAF WADC tech. Rep.*, 1957, No. 56–524.

Miquelon, D. An evaluation of a "regression hypothesis" in skilled motor performance. Unpublished master's thesis, Ohio State University, 1958.

Russell, L. Characteristics of the human as a linear servo-element. Unpublished master's thesis, Massachusetts Institute of Technology, 1951.

Tustin, A. The nature of the operator's response in manual control and its implication for controller design. *J. Inst. elec. Engng.*, 1947, **94**, Part 2a, 190–207.

Walston, C. E., & Warren, C. E. Analysis of the human operator in a closed-loop system. *USAF HRRC res. Bull.*, 1953, No. 53–32.

(Received November 28, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 2, 183–190

THE COMPETENCE OF THE MODEL AND THE LEARNING OF IMITATION AND NONIMITATION

MILTON E. ROSENBAUM AND IRVING F. TUCKER

State University of Iowa

In imitative learning, not only does the learner respond to the cues provided by the behavior of a model but he has available potentially valuable cues in the form of the environmental consequences of the model's response. In their treatment of learning and imitation, Miller and Dollard (1941) consider societal conditions that may facilitate the imitative process, attending especially to the prestige of the models as an important relevant variable. According to their analysis, individuals who have learned to respond independently to environmental cues can, by their behavior, supply relevant cues for the behavior of others for whom the initial environmental cues are obscure, difficult, or not discriminable in some way. Models are regarded as having prestige when the imitation of their behavior consistently leads to a rewarding state of affairs for the imitator.

It is clear that a variety of learned and generalized properties of individuals are valuable clues to prestige, and hence their presence functions to elicit imitative behavior toward a potential model. Miller and Dollard allude to a variety of hierarchies extant in society within which differentiation in prestige is frequently made, i.e., age-grade levels, social status, intelligence, technical knowledge, etc. Underlying many of these dimensions is the availability of knowledge of the environmental competence of a potential model. Competence refers here to the degree to which an individual receives, in reaction to his behavior, evidence of correctness or, more apparently, reward or lack of punishment.

Previous experimental studies of imitation have confounded consistently in their experimental procedures the competence and imitative cues. The model presents his response which is to be matched by the potential imitator and in addition, he is rewarded for his behavior. In such a situation it is not possible to evaluate independently the contribution to the imitative process of the two cues. An exception to this procedural confounding appears only in the work of McDavid (1959) who, when employing adult models for child imitators, did not reward the adults.

By effecting a procedure wherein the imitative cue and the competence cue are independent of each other, the competence cue becomes essentially irrelevant to successful performance of *S*. In the present experiments *S* is capable of learning to imitate or nonimitate solely by attending to the model's response. Differential response in relation to cues of the model's competence may therefore be interpreted in terms of the pre-experimental experiences of *S*s.

It is likely, at least in our cultural milieu, that human organisms learn early to differentiate among others in terms of competence and, accordingly, in terms of prospects for successful imitation. Studies dealing with conformity behavior (Kelman, 1950; Mausner, 1954; Mausner & Bloch, 1957) have indicated the relevance of evidence of competence to the responses of an observer. The present

study investigates experimentally the hypothesis that learning to imitate competent models will occur more readily than learning to imitate incompetent models. A similar analysis may be applied to nonimitative behavior.

In considering the learning of nonimitation, the most apparent and parsimonious hypothesis is simply the obverse of the imitative hypothesis: that an incompetent model's behavior will lead to responses other than imitating and therefore the learning of nonimitation of an incompetent model will obtain a higher level than the learning of nonimitation of a competent model. A variety of other suggestions may be made in predicting nonimitative response, i.e., the irrelevance of past experience with incompetent models, the tendency to nonimitate competent models who in the past received all the available rewards, and the potential salience of competency over incompetent response. The complexity of these possibilities does not permit specific prediction in the present situation.

Experiment I

Method

Apparatus.—The apparatus included two response boxes, two signal boards which presented all of the relevant stimuli to each of 2 *Ss*, and a master control panel which was used by *E*. A response box was placed on the arm of a student chair in each of the two rooms occupied by the *Ss*. It consisted of a double-throw lever switch by means of which each *S* indicated, upon a given signal, a "win" or a "lose" response.

A signal board placed on a table approximately 5 ft. from *S* was also duplicated in each of the two rooms. The signal board consisted of a plywood box 12½ in. wide and 18 in. high, with a depth of 6 in. It was divided into two vertical sections. The top one-third of each of these sections consisted of a compartment containing a 115-v., 7.5-w. frosted lamp, both of which could be turned on and off separately by *E*. The compartments were covered with milk glass: the left compartment was labeled "A" and the right, "B." The lower two-thirds of each vertical section was divided into two compartments, each of which also contained a 115-v., 7.5-w. frosted lamp. The milk glass covering the upper of these compartments on each side was labeled WIN and the lower compartments were similarly labeled, LOSE. All letters were 1 in. high black plastic. Between the upper one-third and the lower two-thirds of each vertical section there was a green and a red 1-in. jeweled lamp.

Procedure.—A simulated horse-racing situation adapted from that used by Brown, Clarke, and Stein (1958) was employed. Two *Ss* participated in each session and each was informed by a sign on the wall of his individual room that he was Subject B. Each *S* was thereby led to believe that the other *S* was Subject A. The *E*, seated in a third room in front of a central control panel, read instructions into a microphone. The instructions indicated that when a signal lamp designating either A or B was illuminated, the appropriate *S* was to guess the outcome of an imaginary horse race. The *S* was to think of the race as involving his horse running against a field of horses, and his task was to guess whether or not his horse would win or lose the race by moving the response lever appropriately. A and B were to think of their horses as running at all times on different tracks in different parts of the country, and as never competing against each other. The *Ss* were also informed that *E* was interested in the group product as well as individual responses. This information was presented in keeping with the finding by Kanareff and Lanzetta (1958) that imitative behavior is facilitated by instructions implying positive sanctions for matching.

The *S*'s guess would be indicated on the signal board in each of the rooms. The accuracy of his guess would then be shown by the jeweled lamps: illumination of the green lamp indicated a correct response; the red lamp, an incorrect response. A single trial included the races of both A and B. Subject A responded first on each trial. Subject A's win or lose response and the red or green lamp remained illuminated during B's portion of the trial.

The similated responses of A were presented by *E*. In every block of 10 trials half of *A*'s responses were win, and half were lose, in a random pattern. The designation of the correctness of the responses of A was made according to programed schedules which

made up the experimental conditions for the differential competence of the model. For one-third of the *Ss*, A was designated correct on 80% of the trials in each block of 10; for one-third A was correct on 50% of the trials; and for one-third A was correct on 20% of the trials.

The designation of the correctness of B's (*S*'s) win or lose response was dependent both on A's response and on the training condition. When imitation was to be trained, a response by *S* which matched that of A was correct. When nonimitation was to be trained, a response by *S* matching that of A was incorrect.

Subjects.—The three levels of the model's competence were presented in conjunction with each of the training procedures: imitation, and nonimitation. Seventy-two male students (12 *Ss* in each of six conditions) were recruited in pairs from the elementary psychology course. The *Ss* were exposed to 10 blocks of 10 trials, and following these, completed a questionnaire adapted from that employed by Kanareff and Lanzetta (1958) which inquired about various aspects of the experiment.

Results and Discussion

The mean frequencies of imitative responses per block of 10 trials for each of the conditions are presented graphically in Fig. 1. It is apparent that the training conditions were effective in producing differential acquisition of imitative responses. Training for imitation yielded consistently increasing frequencies of imitative responses while decreasing frequencies of imitative responses appear under training for nonimitation.

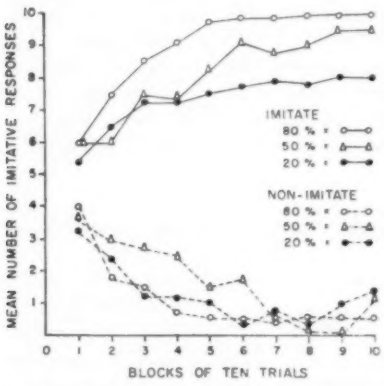

FIG. 1. Mean number of imitative responses per block of 10 trials in Exp. I.

TABLE 1

ANALYSIS OF VARIANCE OF FREQUENCY OF IMITATIVE RESPONSES IN EACH TRAINING CONDITION: EXP. I

Source	*df*	Training Conditions			
		Imitation		Nonimitation	
		MS	*F*	*MS*	*F*
Model's competence (MC)	2	87.85	3.99*	9.70	1.00
Error (B)	33	21.99		14.77	
Trials (T)	9	44.36	23.59**	33.99	16.67**
MC × T	18	2.01	1.07	2.16	1.06
Error (W)	297	1.88		2.04	

* $P < .05$.
** $P < .01$.

Summaries of analyses of variance of the data for each of the training conditions are presented in Table 1. Under training for imitation, the main effects for Model's Competence and for Trials are significant but the Model's Competence × Trials interaction is not. The significant Trials effect reflects the differential operation of the training procedures.

In order to evaluate the experimental hypothesis for learning of imitation in relation to competence of the model the critical difference technique (Lindquist, 1953) was applied to the means presented in the upper row of Table 2. Only the difference between the 80% competence mean and the 20% competence mean is significant. The mean for the 50% competence condition is intermediate. Support is offered for the hypothesis that learning to imitate a competent model proceeds more rapidly than learning to imitate an incompetent model.

The analysis of the data for training to nonimitate yields a significant

TABLE 2

MEAN FREQUENCY OF IMITATIVE RESPONSES IN EXP. I

Training Condition	Model's Competence		
	80%	50%	20%
Imitation	9.07[a]	8.13[ab]	7.36[b]
Nonimitation	1.30[c]	1.78[c]	1.26[c]

Note.—Means in each row marked with a letter in common do not differ significantly at the .05 level by the critical difference technique (Lindquist, 1953).

Trials effect only. This reflects the decreasing tendency to present imitative responses. The nonsignificant main effect for Model's Competence indicates failure to obtain support for the hypothesis that variations in Model's Competence lead to differential patterns in the learning of nonimitation.

Evaluation of the hypotheses was somewhat handicapped by the apparent ceiling imposed by the structure of the task on continued acceleration of learning in both training conditions. A considerable number of *S*s in each of the experimental conditions reached the maximum of either 100% imitation or 100% nonimitation early in the series of trials. Since the procedure allowed *S* in each of the training conditions to be rewarded on every occasion on which *S* responded correctly for his training condition, the frequency of ceiling level performance suggests a rather too simple task. Reduction of the frequency with which *S* can be correct for performing the response that is being trained may reduce the frequency of ceiling level performance and facilitate comparison of the effect of the Model's Competence conditions. Experiment II was conducted in accord with this suggestion.

Before proceeding to Exp. II, data derived from the postexperimental questionnaire will be considered. Ten questions were asked dealing with *S*'s expectations for performance, reports of A's and *S*'s own performance, and reports on *S*'s feelings of frustration and competitiveness. In response to the questions, *S* checked a continuous rating scale ranging from 100% to 0%. Only two of the questions resulted in statistically significant differences among the experimental conditions.

The *S* was asked to report on the model's competence in a question reading, "What percentage of his predictions did your partner predict correctly?" The means for the experimental conditions are presented in Table 3. Analysis of variance indicated a significant main effect only for the Model's Competence variable. The *S*s' reports conform in their pattern to the procedural manipulation of the model's correctness although a tendency toward centrality is apparent. The highly competent model is described as less competent than his actual performance and the reverse is reported for the least competent model.

A question asking for a report on percentage of agreement of responses of *S* and the model resulted in significant differences between *S*s in the different training conditions. The *S*s trained for imitation reported agreement with their partner on an average of 68.6% of the trials while nonimitation *S*s reported an average

TABLE 3

MEAN ESTIMATES OF THE PERCENTAGE OF THE MODEL'S ACCURACY

Training Condition	Model's Competence		
	80%	50%	20%
Imitation	71.3	48.8	37.1
Nonimitation	63.8	50.0	36.3

of 20.75% agreement ($F = 55.12$, $P < .001$). The *Ss* are capable of reporting their behavior during the course of the experiment.

Experiment II

Method

The procedure was identical with that of Exp. I with two modifications. First, for both training conditions, imitation and nonimitation, the maximum frequency of designation of correctness for the prescribed instrumental response was 8 trials of each block of 10 trials. Two trials in each block were chosen at random to be designated as incorrect if the training response was made. Second, because of the possibility that acquisition of the instrumental response would be retarded by the more difficult task, the number of trials was increased from 100 to 150 trials.

The *Ss* were 72 male students (12 *Ss* per condition) in the elementary psychology course, and they served in pairs. Questionnaire data are not available for this experiment.

Results and Discussion

Figure 2 presents the graphical representation of the mean frequencies of imitative responses per block of 10 trials for each condition. It is apparent that the results are similar to those obtained in Exp. I but clearly altered by the imposition of the modified probability of reinforcement. In contrast to the rather smooth acquisition curves of Exp. I, the learning curves for Exp. II suggest considerable block by block variability.

The training conditions appear again to have resulted in differential acquisition of imitative responses. While the curves obtained under training to imitate tend to ascend, those reflecting training to nonimitate tend to descend. Summaries of the analyses of variance for each of the training conditions appear in Table 4.

The results for training to imitate indicate a significant Models × Trials interaction as well as significant main effects for both Models and Trials. The significant interaction term indicates that the three curves differ from each other at least on some of the trials. The critical difference technique applied to the means in the upper row of Table 5 (Lindquist, 1953) indicates that both the 80%

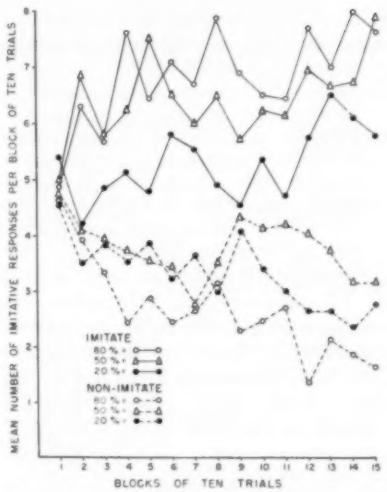

FIG. 2. Mean number of imitative responses per block of 10 trials in Exp. II.

TABLE 4

ANALYSIS OF VARIANCE OF FREQUENCY OF IMITATIVE RESPONSES IN EACH TRAINING CONDITION IN EXP. II

Source	*df*	Training Conditions			
		Imitation		Nonimitation	
		MS	*F*	*MS*	*F*
Model's competence (MC)	2	112.03	5.93**	59.57	4.12*
Error (B)	33	18.88		14.47	
Trials	14	12.34	5.04**	11.98	5.62**
MC × T	28	3.98	1.62*	2.90	1.36
Error (W)	462	2.45		2.13	

* $P < .05$.
** $P < .01$.

TABLE 5

MEAN FREQUENCY OF IMITATIVE RESPONSES IN EXP. II

Training Condition	Model's Competence		
	80%	50%	20%
Imitation	6.84[a]	6.40[a]	5.31[b]
Nonimitation	2.64[c]	3.79[d]	3.36[cd]

Note.—See note to Table 2.

competence and the 50% competence means differ from the mean for the 20% competence condition but do not differ from each other. Further analysis in the form of the Trials × *S*s interactions within each competence treatment group indicates Trials effects ($df = 14/54$) significant at beyond the .01 level for both the 80% ($F = 4.65$) and the 50% ($F = 2.70$) Model's Competence conditions but a nonsignificant effect for the 20% ($F = 1.54$) condition. It appears that learning of imitation occurred under the condition in which the model's competence was high or moderate but not when the competence of the model was low.

The results for training to nonimitate indicate significant main effects both for Model's Competence and for Trials. The interaction of Model's Competence by Trials is not significant. A significant difference is obtained by the critical difference technique (lower row of Table 5) only for the comparison between the 80% and the 50% Model's Competence conditions. Trials × *S*s analysis indicates significant Trials effects for the 80% ($F = 6.17$, $P < .01$) and the 20% ($F = 1.83$, $P < .05$) Model's Competence conditions while the 50% group failed to reach a significant level ($F = 1.47$). Learning of nonimitation was obtained when the model was presented as high or low in competence but not when a moderately competent model was offered.

DISCUSSION

The findings of the present experiments tend to support those of Kanareff and Lanzetta (1958) who demonstrated that the removal of negative sanctions for imitating permits laboratory demonstration of the learning of imitation. In contrast to several previous studies (Bush & Mosteller, 1955; Schein, 1954) imitative learning was accomplished by most *S*s participating in the present studies.

The results indicate that in addition to the specific cues relevant to instrumental response acquisition, information revealing the competence of the model affects imitative learning. In order to perform at a maximum level of instrumental correctness in the present experiment, only the information indicating the response choice of the model need have been attended to by *S*s. As in previous experiments on imitative learning, designation of correctness was contingent on whether *S* matched (or did not match in the case of nonimitation) the response of the model. The correctness of the model's response was essentially irrelevant to successful matching but the results indicate that the variations in model's competence were contributory to the acquisition process. This clearly suggests that *S*'s prior social learning history generalized to the present situation.

As suggested by Miller and Dollard (1941) it is likely that early in life, discriminative learning takes place in which situations, models, drives, and other stimuli are differentiated in terms of their instrumental association with imitative behavior. Although, as with other dimensions, differentiation of models may occur for a variety of attributes, most of these connote some degree of environmental competence. Age, brightness, status, skill, are among the attributes suggested by Miller and Dollard (1941).

The effect of early experiences in the learning of nonimitation may involve several complex features. Although an apparent hypothesis is that individuals learn that incompetent models serve as useful guides to nonmatching behavior, it is also likely that in earlier learning experiences the cues provided by incompetent behavior are primarily irrelevant to successful responses. Competent models, on the other hand, in some situations may present rather reliable cues for nonmatching. This will be true especially in situations wherein the model's competence enables him to claim all the available rewards, leaving the observer only the alternative of a nonmatching response. Alternatively, competence may be consistently highlighted during early experiences so that all cues associated with competence are more salient than those related to incompetence. While Exp. I presents no differences among Model's Competence conditions in learning to nonimitate, the data of Exp. II indicate facilitation of learning of nonimitation in the 80% Model's Competence condition in comparison to the 50% condition. A trend in the direction of facilitation of acquisition of the nonimitative response appears for the 20% Model's Competence group but statistical significance is not reached. The results for this group may reflect the conflicting response tendencies based on the competing early experiences with incompetent models that were suggested above.

Kanareff and Lanzetta (1958), Lanzetta & Kanareff (1959), and Niemark and Rosenberg (1959) found that imitative response did not achieve correspondence with the probability of designation of correctness to which *Ss* were exposed. Although in each of the present experiments nonvarying probabilities were employed (100% in Exp. I and 80% in Exp. II), in each case the matching of the probabilities of correctness and imitative response was quite close for several of the experimental groups. In Exp. I under training to imitate, *Ss* in the 80% Model's Competence condition imitated during the last five blocks of trials at a level continuously exceeding 98%. On only six of the trials was there failure to match on the part of any *Ss* in this group. All but 6 *Ss* trained to nonimitate achieved levels of response of nonmatching approximating the probability of correctness. The 6 *Ss* were distributed over the three Model's Competence conditions.

In Exp. II coincidence of the instrumental response and probability of correctness was approximated only by *Ss* in the 80% Model's Competence conditions. The mean of the last five blocks of trials for those in this condition who were trained to imitate was 7.37, while the corresponding mean for nonimitators was 1.91.

Another finding that is in contrast to results obtained by Kanareff and Lanzetta (1958), Lanzetta & Kanareff (1959), and Niemark and Rosenberg (1959) concerns differential rapidity of learning of imitation and nonimitation. Their results indicate more rapid acquisition of imitation than nonimitation. These findings are not repeated in the present data and, in fact, a trend in the direction of more rapid acquisition of the instrumental response is indicated for the nonimitation condition in the two experiments (cf. Fig. 1 and 2). This is consistent with Bush and Mosteller's (1955) suggestion that events which increase nonimitation have a greater effect than events that increase imitation.

The present experiments differ in two important ways from previous studies in which conformity behavior or matching has been examined in relation to the competence of the model (Kelman, 1959; Mausner, 1954; Mausner & Bloch, 1957). In the present situation, the course of acquisition of the dependent response was examined in contrast to simple measurement of its current frequency of appearance. In addition, information as to the model's competence was revealed sequentially in the context of serial response rather than as a pre-experimental manipulation.

Summary

The effect of information concerning the competence of a model on learning of imitative and nonimitative behavior was examined in two experiments. The *Ss* predicted the outcomes of a series of fictitious horse races after exposure on each trial to the prediction and correctness of the prediction made by a simulated partner. In three conditions varying the model's competence, the partner was correct on 80%, 50%, or 20% of each block of 10 trials.

In Exp. I, half of the *Ss* were trained under conditions wherein they were correct 100% of the time if their predictions matched the response of the partner. The other half of the *Ss* were correct on 100% of the trials if they nonmatched the partner's prediction. The procedure allowed independent designation of correctness by specifying that the horse races predicted in each trial were run on separate race tracks. Ten blocks of 10 trials were employed. Under training to imitate, the results indicate that the greater the model's competence, the greater is the facilitation of the learning process. Under training to nonimitate, no differences appeared among the three conditions varying the model's competence. Questionnaire data indicated that *Ss* were capable of accurate report of several aspects of the procedure.

Experiment II was designed to avoid a ceiling effect by allowing *Ss* to be correct a maximum of 80% of the trials, and the number of trials was extended to 150. The results for training to imitate again indicated that the greater the model's competence, the greater is the facilitation of acquisition of the imitative response. Under training to nonimitate, learning was facilitated when the model's competence was high or low but was poorer when the model's adequacy was mediocre.

The results were discussed as reflecting the effect of prior social learning experiences on current social behavior.

References

Brown, J. S., Clarke, F. S., & Stein, L. A new technique for studying spatial generalization with voluntary responses. *J. exp. Psychol.*, 1958, **55**, 359–362.

Bush, R. R., & Mosteller, F. *Stochastic models for learning.* New York: Wiley, 1955.

Kanareff, V. T., & Lanzetta, J. T. The acquisition of imitative and opposition responses under two conditions of instruction-induced set. *J. exp. Psychol.*, 1958, **56**, 516–528.

Kelman, H. C. Effects of success and failure on "suggestibility" in the autokinetic situation. *J. abnorm. soc. Psychol.*, 1950, **45**, 267–285.

Lanzetta, J. T., & Kanareff, V. T. The effects of a monetary reward on the acquisition of an imitative response. *J. abnorm. soc. Psychol.*, 1959, **59**, 120–127.

Lindquist, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1953.

McDavid, J. W. Imitative behavior in preschool children. *Psychol. Monogr.*, 1959, **73**(16, Whole No. 486).

Mausner, B. The effect of prior reinforcement on the interaction of observer pairs. *J. abnorm. soc. Psychol.*, 1954, **49**, 65–68.

Mausner, B., & Bloch, B. L. A study of the additivity of variables affecting social interaction. *J. abnorm. soc. Psychol.*, 1957, **54**, 250–256.

Miller, N. E., & Dollard, J. *Social learning and imitation.* New Haven: Yale Univer. Press, 1941.

Neimark, E. D., & Rosenberg, S. The effect of "social" discriminative cues on probability learning. *J. exp. Psychol.*, 1959, **58**, 302–311.

Schein, E. H. The effect of reward on adult imitative behavior. *J. abnorm. soc. Psychol.*, 1954, **49**, 389–395.

(Received November 28, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 2, 191–197

EFFECTS OF MENTAL SET UPON VESTIBULAR NYSTAGMUS[1]

WILLIAM E. COLLINS

Civil Aeromedical Research Institute, Federal Aviation Administration, Oklahoma City

During a course of repeated rotational stimulation, vestibular nystagmus has been reported to undergo "habituation," i.e., the oculomotor reaction and, in man, the subjective sensation of rotation, decline (Aschan, 1954; DeWit, 1953; Dodge, 1923; Griffith, 1920; Guedry & Ceran, 1959). Wendt (1951), however, holds that such declines occur only when visual stimulation is permitted or when, in total darkness, *S* does not maintain an "environment-directed orientation." He says ". . . (in darkness) loss of nystagmus occurs when the subject fails to remain alert to his external surroundings. This habituation is preventable by keeping the subject attentive to the external surrounds and avoiding inward-directed revery states" (p. 1214).

To evaluate such a mental set variable, a study was conducted (Collins, Crampton, & Posner, 1961) to determine the effect of instructions, which would influence *S*'s mental activity, on the nystagmic response elicited by angular acceleration. Significantly less nystagmus was produced during stimulation when *S*s were requested to assume reverie states than when they were kept alert throughout a trial either by signaling subjective events or by doing prolonged arithmetic problems, although there were certain inconsistencies in the results. The design of this first study was such as to place considerable demands upon *S*s, i.e., (*a*) they were required to shift rapidly from mental activity to reverie states in successive trials, and (*b*) the period between instructions to relax and daydream and the start of the rotational stimulus was quite brief (30 sec.). It is likely that these factors account for the inconsistencies noted.

The present study was undertaken to clarify the effects of various types of mental states on vestibular nystagmus by establishing favorable conditions for obtaining specific sets. Tasks and instructions were designed to produce four types of subjective conditions: (*a*) mentally alert, attending to the rotational stimulus, and signaling subjective events; (*b*) mentally alert and attending to an environmental stimulus other than that of rotation; (*c*) mentally alert and ignoring the environment; (*d*) in a state of reverie and ignoring the environment.

METHOD

The stimulator, situated in a light-proof room, consisted of a turntable on which *S* was seated with his head in the center of rotation. The rotary apparatus is described in detail elsewhere (Guedry & Kalter, 1956) and was the same as that used in the previous report (Collins et al., 1961). The modifications introduced here were the employment of a buzzer and a bell located directly behind *S*, a Hunter timer used by *E* to control the duration of the buzzer, and a telegraph key manipulated by *S* and connected to both the buzzer and the recording device. A second

[1] The author is indebted to F. E. Guedry, Jr., G. H. Crampton, and L. S. Caldwell for their many helpful suggestions, and to Kenneth J. Gall for his excellent technical assistance.

This study was conducted at the United States Army Medical Research Laboratory, Fort Knox, Kentucky.

telegraph key, used to signal subjective events, was also connected to the recorder.

Two criteria were used in the selection of *S*s: (*a*) no history of labyrinthine disorder; (*b*) no previous experience in vestibular experimentation. Three electrodes were employed with each *S*. Two of these provided measurement of horizontal eye movements and were taped on the temporal side of each eye, slightly to the rear of the junction of the temporal and zygomatic bones. The third electrode was taped behind the left ear on the mastoid process.

The amplifying and recording device was an Offner Type T electroencephalograph. Three channels were used. One of these recorded the horizontal component of vestibular nystagmus, another denoted the periods of acceleration and deceleration, and the final one recorded signals made by *S* with either of the two telegraph keys.

Prior to each trial, 20° eye-movement calibrations were obtained from each *S* and used as conversion factors for the later data measurements. All trials were conducted in complete darkness with rotation in a clockwise direction. The *S*'s eyes were open and his head held upright by means of a bite-board.

Four conditions were employed: (*a*) Key pressing. The *S* was required to signal the onset, cessation, and each experience of 90° of rotation. This technique, described by Guedry and Lauver (1961), keeps *S* attending to the sensation of rotation. (*b*) Mental arithmetic. The *S*s were instructed to ignore the rotational environment and to do mental arithmetic problems as rapidly, as continuously, and as accurately as possible. They were given a constant divisor and a starting dividend and were required to divide each successive quotient by that divisor. Answers were recorded at the conclusion of each trial. The task was designed to provide a means of keeping *S* "inwardly directed" but mentally alert and active. (*c*) Temporal reproduction. The *S*s were instructed to attend carefully to various durations of sound stimuli (buzzer) presented at frequent intervals throughout each trial. They were to listen carefully for the start of each such stimulus and, at its conclusion, reproduce the duration of the sound by means of one of the telegraph keys. Pressing this key both activated the buzzer and marked the duration of the response on the recording paper. This condition kept *S* "environmentally directed," but with his attention oriented toward time estimation rather than the experience of rotation. (*d*) Reverie. The *S*s were asked to relax as completely as possible, but to keep their eyes open. They were told that they could daydream, but were asked not to think about anything of real concern to them, nor deliberately to pursue any line of thought.

Each *S* was tested for 5 successive days. On Day 1, four practice trials, one under each mental condition, were given. The following four daily experimental sessions each consisted of six trials with rest periods of 3–4 min. interspersed between them. The same condition was employed throughout a session for a given *S*, but the conditions differed from day to day. Further, a counterbalancing of conditions was employed with 2 *S*s assigned to each of the four established orders of presentation.

The procedure for the experimental trials consisted of the following: (*a*) 1 min. of rotation at a constant velocity of 1 rpm, (*b*) 15 sec. of acceleration at 4.5°/sec², (*c*) 2 min. at a constant velocity of 12.2 rpm, (*d*) 15 sec. of deceleration at 4.5°/sec², and (*e*) 2 min. at a constant velocity of 1 rpm.

The key pressing and mental arithmetic conditions were further characterized by the sounding of a bell for a period of 2 sec., 12 sec. prior to acceleration. This served to signal *S* to begin work on the arithmetic problem in the one case, and to alert him to the impending start of rotational stimulation in the key pressing situation. The bell was also sounded 12 sec. prior to deceleration in the later condition for the same reason. The sound stimuli for reproduction of duration were interspersed at definite points throughout each trial. The stimuli ranged from 0.5 to 6.5 sec. at 1-sec. intervals. Each of these values was presented twice in a given trial.

Results

The recordings were analyzed for two types of data: nystagmus output (degrees of slow-phase eye movement), and duration of response. These data were scored separately for acceleration and deceleration. The duration of the vestibular reaction was measured in seconds from the first nystagmic beat to the end of the primary response. The output data were scored in centimeters with a Keuffel and Esser map measure (Model 1741) and transformed to

TABLE 1

THE AVERAGE OUTPUT AND DURATION OF NYSTAGMIC RESPONSE FOR EACH OF THE FOUR EXPERIMENTAL CONDITIONS

Ss	Nystagmus Output (Degrees)				Nystagmus Duration (Seconds)			
	Mental Arith.	Key Press	Temporal Reprod.	Reverie	Mental Arith.	Key Press	Temporal Reprod.	Reverie
	Acceleration							
1	616	684	402	384	57	49	41	43
2	505	457	320	55	39	35	32	7
3	708	561	531	364	56	43	41	33
4	361	452	189	343	40	41	24	33
5	609	509	580	510	55	49	50	47
6	930	824	750	418	51	44	46	33
7	367	225	329	44	37	31	37	17
8	408	329	218	39	31	33	27	4
Mean	563	505	415	270	46	41	37	27
SD	195	190	193	192	10	7	9	16
	Deceleration							
1	524	319	329	253	50	35	36	38
2	317	259	160	5	38	31	21	1
3	549	619	615	51	39	41	48	18
4	451	455	422	438	51	44	38	36
5	626	565	622	407	74	62	64	47
6	896	795	721	439	42	40	39	30
7	635	328	421	53	47	32	38	20
8	485	435	336	73	37	36	28	11
Mean	560	472	453	215	47	40	39	25
SD	169	180	187	191	12	10	13	15

TABLE 2

ANALYSES OF VARIANCE FOR THE NYSTAGMUS OUTPUT AND THE NYSTAGMUS DURATION DATA

Source	*df*	Error Term	Nystagmus Output		Nystagmus Duration	
			MS	*F*	*MS*	*F*
(1) Subjects (*S*)	7	4	196,765	7.91***	809	7.13***
(2) Direction (D)	1	4	2,773	0.11	0	0.00
(3) Conditions (C)	3	5	298,311	20.85***	1,159	21.19***
(4) *S* × D	7	7	24,884	4.76**	113	5.77**
(5) *S* × C	21	7	14,305	2.73*	54	2.78*
(6) D × C	3	7	6,563	1.25	14	0.73
(7) *S* × D × C	21	—	5,231	—	19	—
Total	63					

* $P < .05$.
** $P < .01$.
*** $P < .001$.

degrees of eye movement by means of the conversion factors. The measurements made represent the vertical distances from the peak of each nystagmic beat to the slow-phase baseline of that beat, accumulated for the 30-sec. period beginning with the acceleration or deceleration stimulus.

Mean values for each *S*, based on the six trials for each experimental

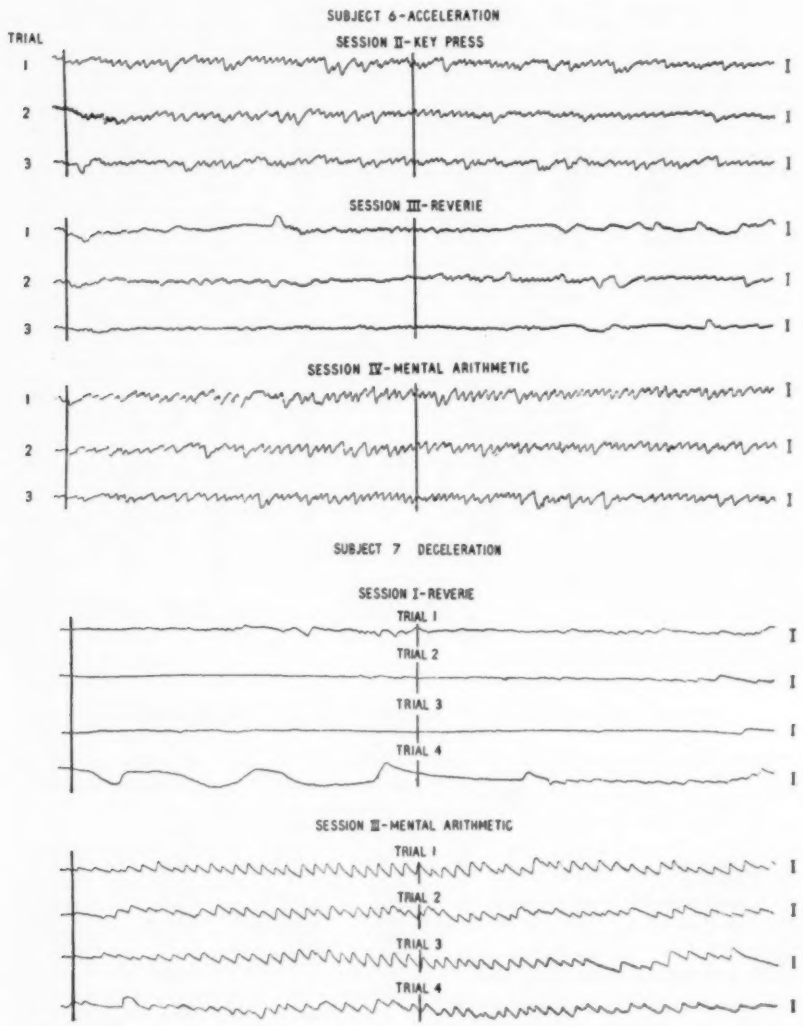

FIG. 1. The recorded nystagmus elicited by acceleration and deceleration stimuli for 2 *S*s. (Vertical bars through each record denote the 15-sec. periods of stimulation and calibration markers on the right indicate 20° eye excursions.)

condition, are presented in Table 1. These data were used in the separate analyses of variance conducted for each of the two nystagmus measures (Table 2).

The analyses and *t* tests for repeated measurements on the same *S*s yielded an identical pattern of results for both the output and the duration of nystagmus data. In neither case was the direction of stimulation (acceleration and deceleration) a significant variable. Significant *F* ratios were obtained for *S*s and for the four conditions of mental set. Since there were no directional differences, *t* tests were conducted among the conditions using pooled acceleration and deceleration scores. For both output and duration, the results of the *t* tests indicated: (*a*) significantly less nystagmus during reverie states than during any of the three attention-demanding tasks; and (*b*) significantly more nystagmus for the mental arithmetic condition than for any of the other tasks.

Some representative examples of the recorded nystagmic responses appear in Fig. 1. Both acceleration and deceleration data are presented. The marked reduction in nystagmus during reverie conditions, as well as the sharply defined, regular beats that occur during the mental arithmetic task, are readily apparent.

As with the main effects, a like pattern of significant interactions appeared from the two analyses of variance. In both cases, significant *F*s were obtained for $S \times$ Conditions and for $S \times$ Direction of Stimulation. These relationships indicate that the stimulus direction and the mental set conditions exert some differential effects upon different *S*s, e.g., although the overall effects of stimulus direction were not significant, some *S*s produced less nystagmus as a result of the deceleration stimulus, while others produced more, when compared to the output and duration of response to acceleration.

TABLE 3

RESULTS OF THE *t* TESTS AMONG THE EXPERIMENTAL CONDITIONS FOR THE NYSTAGMUS OUTPUT AND THE NYSTAGMUS DURATION DATA

Conditions	Mental Arith.	Key Press	Temporal Reprod.	Reverie
			Nystagmus Output	
Mental Arith.		2.74*	5.49***	5.50***
Key Press	4.27**		1.65	5.39***
Temporal Reprod.	4.89**	1.10		3.54**
Reverie	7.32***	4.26***	3.32*	
	Nystagmus Duration			

* $P < .05$.
** $P < .01$.
*** $P < .001$.

DISCUSSION

Nystagmus.—The data obtained by instructing *S*s to assume different states of attention point clearly to the importance of the subject's mental set in the production of vestibular nystagmus. As Wendt (1951) indicated, reverie states result in a decline of the nystagmic response, but the statement that the response will be maintained from trial to trial as long as *S* remains "environment-directed" requires qualification. If this statement is interpreted to mean that *S* must attend either to the sensations aroused by rotation or to other environmental stimuli, then the key pressing and temporal reproduction conditions, respectively, fulfill such requirements and do provide means for maintaining a high level of nystagmic output. However, mental arithmetic is at least equally effective in this regard even though the subjects are frequently unaware of the external environment, including the ro-

tational stimulus, while working on the arithmetic problems.

What appears to make the arithmetic task so effective is that it requires a continuous and concerted state of alertness on the part of *S*. Further, at least with *S*s used here, the task assumes the nature of a challenge. The *S*s expressed concern about their accuracy and speed compared with others.

The estimates of subjective velocity made through the key pressing technique, while keeping *S*s alert, appear to have one drawback, i.e., habituation. The experience of rotation has been shown to decline with repeated stimulation (Guedry & Ceran, 1959). Under such conditions, as the sensation decays, *S*s probably tend to relax thereby reducing both the magnitude and duration of the nystagmic reaction (Guedry, Collins, & Sheffy, 1961; Guedry & Lauver, 1961).

Reproduction of sound durations ranked third in efficiency of maintaining the vestibular response. It is possible that its power could be increased. In the present study the sound level of the buzzer was quite high—such that *S*s had no difficulty in hearing the stimuli. Thus it was not necessary that *S* studiously attend to the stimulus lest he fail to hear it. It is felt that this method of maintaining alertness might be improved by using sound stimuli that are close to threshold, thereby requiring higher levels of attention on the part of *S*s.

One of the main factors, then, in the maintenance of vestibular nystagmus, would seem to be the state of alertness of *S* during stimulation. Such an explanation accounts for the differences obtained by Mowrer (1934) with "rested" and "excited" pigeons, and for other findings (Griffith, 1920) that demonstrated an enchanced vestibular response immediately after *S*s had engaged in strenuous physical exercise. In both cases, the animals and men were "keyed up" when tested. Further, the absence of a "constant reaction" noted for some *S*s in a recent study by Suzuki and Totsuka (1960) is probably attributable to fluctuations in alertness. That a "temporary reduction in response followed by a complete and lasting restoration of nystagmus" (p. 575) also occurred seems to indicate that the states of alertness of *S*s were somehow changed, perhaps by questions or by instructions.

These findings also agree with a point made previously in an investigation concerned with caloric stimulation of the labyrinth (Mahoney, Harlan, & Bickford, 1957). It was noted that when *S*s closed their eyes during a caloric nystagmus, eye movements were often completely inhibited. However, performance by *S*s of a short arithmetic task at this time "frequently elicited a brief burst of nystagmus" (p. 52).

The usual conditions under which rotational stimulation of the vestibular system is experimentally applied in total darkness tend to favor the development of a reverie attitude. However, that such states account for all of the response decline during a course of repeated rotation has not been conclusively demonstrated for man. In this regard, it should be noted that, for cat, a marked loss of vestibular nystagmus could be recovered only partially by such alerting techniques as electric shock and high intensity auditory stimulation (Crampton & Schwam, 1961). The use of d-amphetamine does not prevent this decline (Crampton, 1961).

Individual differences.—Although it is clear that reverie can strongly influence vestibular nystagmus, there are wide individual differences in the effect of this mental state. For some *S*s the nystagmus decline may be so great that the response appears to be completely absent, or the amplitude of the beats, when they occur, is so small as to preclude the possibility of scoring by the technique employed here. And yet, for 1 *S* (*S*4; see Table 1), the decline during reverie is negligible.

All *S*s reported that they had little or no difficulty in relaxing and daydreaming. However, this may mean different things, both physiologically and psychologically, to individual *S*s. Thus, it would seem that some type of objective task should be assigned to *S*s

employed in vestibular research in order to assure, to the highest possible degree, uniformity of mental set.

Summary

During rotational stimulation *Ss* were instructed, in different test sessions, to: (*a*) do mental arithmetic, (*b*) make estimates of subjective velocity, (*c*) reproduce durations of sound stimuli, and (*d*) assume a state of reverie. The reverie state resulted in significantly less recorded nystagmus than the other three tasks, while mental arithmetic was the most efficacious means of obtaining a vigorous response. A factor of sustained alertness seems to be an important condition for the maintenance of nystagmus.

References

Aschan, G. Response to rotary stimuli in fighter pilots. *Acta oto-laryngol., Stockh.*, 1954, Suppl. No. 116.

Collins, W. E., Crampton, G. H., & Posner, J. B. Effects of mental activity on vestibular nystagmus and the electroencephalogram. *Nature, Lond.*, 1961, **190**, 194–195.

Crampton, G. H. Habituation of vestibular nystagmus in the cat during sustained alertness produced by d-amphetamine. USA Med. Res. Lab. Rep. 1961, No. 488.

Crampton, G. H., & Schwam, W. J. Effects of the arousal reaction on nystagmus habituation in cat. *Amer. J. Physiol.*, 1961, **200**, 29–33.

DeWit, G. Seasickness: (Motion sickness) a labyrinthological study. *Acta oto-laryngol., Stockh.*, 1953, Suppl. No. 108.

Dodge, R. Habituation to rotation. *J. exp. Psychol.*, 1923, **6**, 1–35.

Griffith, C. R. The organic effects of repeated bodily rotation. *J. exp. Psychol.*, 1920, **3**, 15–46.

Guedry, F. E., Jr., & Ceran, S. J. Derivation of "subjective velocity" from angular-displacement estimates made during prolonged angular accelerations: Adaptation effects. USA Med. Res. Lab. Rep., 1959, No. 376.

Guedry, F. E., Jr., Collins, W. E., & Sheffy, P. L. Perceptual and oculomotor reactions to interacting visual and vestibular stimulation. *Percept. mot. Skills*, 1961, **12**, 307–324.

Guedry, F. E., Jr., & Kalter, H. Description of human rotation device. USA Med. Res. Lab. Rep., 1956, No. 242.

Guedry, F. E., Jr., & Lauver, L. S. The oculomotor and subjective aspects of the vestibular reaction during prolonged constant angular acceleration. *J. appl. Physiol.*, 1961, **16**, 215–220.

Mahoney, J. L., Harlan, W. L., & Bickford, R. G. Visual and other factors influencing caloric nystagmus in normal subjects. *Arch. oto-laryngol.*, 1957, **66**, 46–53.

Mowrer, O. H. Influence of "excitement" on the duration of post-rotational nystagmus. *Arch. oto-laryngol.*, 1934, **19**, 46–54.

Suzuki, J., & Totsuka, G. Post-rotatory nystagmus: Modifications observed in experiments with repeated rotatory stimulation. *Acta oto-laryngol., Stockh.*, 1960, **51**, 570–578.

Wendt, G. R. Vestibular functions. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 1191–1223.

(Received November 30, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 2, 198–206

RESPONSE TO SIMULTANEOUS STIMULATION OF TWO SENSE MODALITIES[1]

JACK A. ADAMS AND RIDGELY W. CHAMBERS[2]

University of Illinois

Is it possible for a human *S* to do two things at once without impairment? Stated more concretely, if *S* is presented a light and a tone simultaneously, and must respond to the light with one hand and to the tone with the other, will his response measures to the light (or the tone) be the same as when each stimulus is presented separately for one-hand response in a simple, unisensory task? Basically, the foregoing statement expresses the design of the experiment reported here, and it is the one considered sufficient to examine the question. Viewed superficially, no impairment might be expected for the compound task because each stimulus source with its own sense modality and each response with its own effector system can be considered a separate, independent S-R loop. But, considered critically, the likelihood of sensory interaction for two simultaneous stimuli, central interaction, or interaction from the motor feedback of two responses, suggest reasons why performance in the complex task might be different from that of unisensory control tasks (Adams, 1961).

The present experiment was designed to be a careful laboratory test of the fundamental load-carrying capabilities of the human *S*. A discrete bisensory tracking task was used in either a unisensory version or a bisensory version where a stochastic series of simultaneous stimulus pairs was presented, one audio and the other visual, each for response by a separate hand. The scope of the investigation was extended by having some pairs perfectly predictable with learning (event certainty), and some unpredictable (event uncertainty). All events had constant duration and thus time certainty. It was hypothesized that the prediction of an event pair, both in type and time, could often free *S* of directly sensing environmental stimuli before responding and this would minimize sensory interaction as a possible source of impairment.

METHOD

Apparatus.—The research device was a version of the two-dimensional, all-visual, discrete tracking apparatus described by Adams & Xhignesse (1960), but modified to also allow two-dimensional bisensory tracking and one-dimensional tracking to either visual or auditory stimuli. In the visual dimension, *S* had three horizontally-arranged jeweled stimulus lights in front of him at eye level. The lights were red, white, and green, with white in the middle. Each light was .5 in. in diameter, and the three lights were mounted in a $2 \times 3\frac{1}{4}$ in. box. Beneath each stimulus light was a small neon response feedback light that informed *S* of the position of his control, and this direct display of response feedback cues, as well as stimulus input lights, made the task one of pursuit tracking. The stimulus lights came on in a probabilistically-defined sequence and were programed on punched tape and automatically read by a motorized tape reader, Model 2, manufactured by the Commercial Controls Corporation. The *S* had to keep moving the control so the feedback light was under the lighted stimulus as much as possible.

[1] This research was supported by the United States Air Force under Contract No. AF 49(638)-371, monitored by the Air Force Office of Scientific Research of the Air Research and Development Command.

[2] Now at the University of Arizona.

The auditory dimension was designed to have the same pursuit tracking properties as the visual dimension. There was a low, a medium, and a high pure tone (600, 800, and 1000 cps), and they came on in a probabilistically-defined order. The *S* heard them over a headset and responded with the same control as used for visual stimuli. The auditory error coding was in terms of a complex tone. When *S* had the control in the correct position he heard the pure tone but, when he was wrong, he heard a complex tone made up of two of the three fundamental frequencies. The correct stimulus was presented as a pure tone in the complex and, superimposed on it was a second tone, rapidly interrupted, whose frequency was determined by the position of the control. Thus, when *S* was in error he was provided direct feedback on the correct stimulus and the control position, and this was formally equivalent to the visual pursuit tracking task.

Response was with a 2-in. control stick mounted on a wide armrest of *S*'s chair, and *S* could have a control for one or for each hand, depending on whether he was in a unisensory or bisensory group. Either sensory dimension could be assigned for response by either hand, and angular orientation of the control could be set by *E* because the entire control assembly could be rotated and locked for any direction of control movement. The control made electrical contact by way of a small extension arm under the armrest of the chair, and the arm had a ballbearing tip that could be rolled to any one of three small contact plates set in a plastic hemispheric disk. The distance between the three contacts was negligible and this design feature, coupled with appropriate electrical circuitry, gave no dead space between control positions so *S* was always on one of three contacts and always had knowledge of control position because one of the feedback lights was always on.

Performance measurement was by six time-on-target timers, one for each of the three stimulus events in each sensory dimension. An Esterline-Angus operations recorder allowed *E* to record the time relationships between all stimulus and response events on a trial.

Experimental procedure.—Group A was a unisensory control group and responded only to auditory stimuli. Group V was the other control group and received only visual stimuli. Group AV was the bisensory experimental group where a visual and an auditory stimulus of the same duration were always presented simultaneously, the visual event for response by one hand and the auditory event for response by the other. Using a compass reference, and considering a backward-forward control movement at a right angle to *S* as 0°, a control's directional orientation was either 45° or 315°. For Group AV, both controls had either divergent or convergent lines of movement. Within each group, the direction of control movement and right or left hand was counterbalanced among *S*s. Group AV had further counterbalancing of auditory and visual stimuli with respect to control direction and hands.

Each trial was 100 sec. in duration and consisted of 50 2-sec. events or, in the case of Group AV, 50 pairs of 2-sec events. Sixty practice trials were given. There were three practice sessions, each on a different day, of 20 trials each. Intertrial rest was 40 sec., except between Trials 10 and 11 when it was 5 min. All events and rest periods were programed on punched tape and proceeded without intervention from *E*. Time on target was recorded on every trial, but the Esterline-Angus operations recorder was used only on Trials 56 and 57 to provide for a detailed analysis at a high level of practice.

Characteristics of stimulus series.—Table 1 gives the theoretical probabilities used in generating the sequence of intramodality stimulus events in both auditory and visual, as well as the conditional probabilities for intermodality event pairs showing the probability of a visual event being paired simultaneously with an auditory event. To illustrate the implications of Table 1 for the predictability of events, it will be noted that conditional probabilities for intramodality stimuli have both certain events (white light or medium tone) and uncertain events (red or green light, high or low tone). A series of visual events could look like this: red, white, red, white, green, white, red, white, green, etc. A white event followed a colored event with certainty, but a white event had an equal chance of being followed by a red or green event. The auditory events had the same patterning. When both sensory dimensions were used for Group AV, there was the added requirement for specifying the probability relationship between the auditory and visual events that were always paired and presented simultaneously, and the implication of the intermodality conditional probabilities is that the certain white event and the certain medium tone always occurred together, and the uncertain red and green lights each had

TABLE 1

THEORETICAL PROBABILITIES OF OCCURRENCE, $p(i)$, AND CONDITIONAL PROBABILITIES, $p_i(j)$ AND $p_k(l)$, FOR EVENTS OF THE VISUAL AND AUDITORY SERIES

Probability of Occurrence			Conditional Probabilities for Intermodality Event Pairs			
Visual Stimulus (Light)	Auditory Stimulus (Tone)	$p(i)$	k	l		
			$p_k(l)$	High	Medium	Low
Red	High	.25	Red	.50	0	.50
White	Medium	.50	White	0	1.0	0
Green	Low	.25	Green	.50	0	.50

Conditional Probabilities for Intramodality Events

Visual Events				Auditory Events			
i	j			i	j		
$p_i(j)$	Red	White	Green	$p_i(j)$	High	Medium	Low
Red	0	1.0	0	High	0	1.0	0
White	.50	0	.50	Medium	.50	0	.50
Green	0	1.0	0	Low	0	1.0	0

Note.—The notations i and j refer to events within a particular sense modality, and $p(i)$ and $p_i(j)$ define the organization of events for both the visual and auditory series. The notations k and l refer to two simultaneous events, one auditory and one visual, and $p_k(l)$ is the conditional probability that a kth visual event will be paired with an lth auditory event.

an equal likelihood of being paired with the uncertain high and low tones.

Both the visual and the auditory series had a redundancy of 68% (Attneave, 1959). The procedure was to generate one tape of 1000 pairs of events, and this same tape was used for each group on each of the three sessions. Group AV was given all events on the tape, but one of the stimulus sources was withdrawn for Group A or Group V. A response to a certain event was the center position of the three on the control, and responses to uncertain events were on either side of the center.

Subjects.—There were 16 *S*s in a group. The 48 *S*s were university male undergraduates who were paid for their participation. They were randomly assigned to groups.

RESULTS

Method of data summary and presentation.—The stimulus events for the three groups represent eight experimental conditions, depending on whether the modality was visual or auditory, the events certain or uncertain, and whether the event was presented alone or simultaneously with another event in the other sense modality. For each modality there were two types of uncertain events—red and green lights for the visual, and high and low tones for the audio. The measures for the two uncertain events of a modality were pooled as a common class of uncertain events because their probabilities of occurrence were the same and they posed the same response demands for *S*. The equipment provided performance measures for each of the eight conditions. Each response measure is designated by its stimulus class, and is always labeled either C or U for certain and uncertain, and A or V for auditory or visual. In addition, the measures of Group AV will have a lower case a or v to signify that a stimulus occurred simultaneously in the other sense modality. For example, a CA measure is the response of Group A to certain auditory events,

and CAv is the same response measure for Group AV, but where each certain auditory event is simultaneously accompanied by a certain visual event that is being given response with the other hand.

Time on target.—Figure 1 shows the performance curves of the eight conditions. Time on target for certain events is higher than for uncertain events in all cases (Bricker, 1955). Four Groups × Trials analyses of variance (Type I mixed design—Lindquist, 1953) were performed on the final 10 trials to determine differences between single and bisensory responding. The comparisons of principal interest were between groups, and Table 2 shows that time on target for uncertain bisensory events was significantly poorer than for uncertain unisensory events, and this held for both auditory and visual stimuli. But this was not the case for certain events because Cond. CaV had a significantly *higher* performance level than CV. This superiority for bisensory certain events was true only for the visual modality because the difference in performance between Cond. CA and CAv lacked the .05 level.

Response time analysis.—Time-on-target proficiency in discrete tracking is a function of (*a*) time between the onset of the stimulus and the onset of the response, whether the response is correct or incorrect, (*b*) number of errors, i.e., movements of the control to the wrong position, and (*c*) duration of each error before it is corrected. The operations recorder on Trials 56 and 57 allowed measurement of each of these response determiners of tracking proficiency, and tabulations were made of error frequency, and the time between onset of the stimulus and the onset of the response,

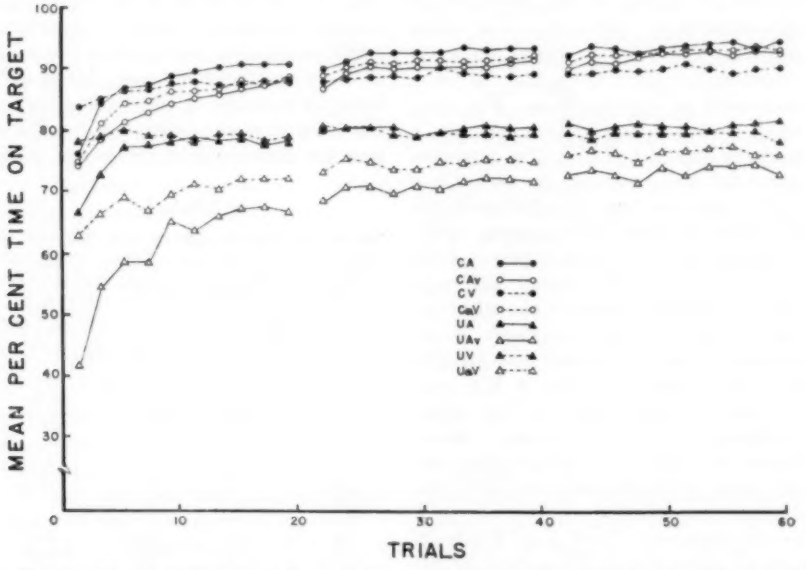

FIG. 1. Mean percentage of time on target on blocks of two trials for the eight experimental conditions. (The breaks in the curves after Trials 20 and 40 signify intersession intervals.)

TABLE 2

F RATIOS OF FOUR GROUPS × TRIALS ANALYSES OF VARIANCE OF THE TIME-ON-TARGET SCORES OF TRIALS 51–60

Source of Variance	*F* Ratio			
	CV & CaV	CA & CAv	UV & UaV	UA & UAv
Between groups	5.30*	0.60	6.28*	49.02**
Trials	1.33	2.03*	5.36**	2.58*
Groups × Trials	1.26	1.65	2.04*	1.04

* $.01 < P < .05$.
** $P < .01$.

whether the response was correct, incorrect, or correction of an error. These measures are called *response times* (RT) to distinguish them from reaction time measures where steps are taken to insure that response follows the stimulus. In a discrete tracking task a response can either follow or be anticipatory and occur before the onset of the stimulus or simultaneously with it. The time between each stimulus event and each response on Trials 56 and 57 was measured in milliseconds. The convention was followed of assigning a positive sign to RT if the response preceded the onset of the stimulus, and a negative sign if it followed the stimulus (Adams & Xhignesse, 1960; Poulton, 1952). Three types of RTs were recorded. A Correct RT was the time between onset of the stimulus and onset of the response when response was the correct one for the stimulus presented. Error RT was measured in the same way but when a response was initially wrong. Error Correction RT was the time between the onset of the stimulus and the time taken to complete the correction of a wrong response. Each *S* on Trials 56 and 57 contributed RT values for 50 certain and 50 uncertain events to the analysis. The algebraic mean in milliseconds was taken of the values of each of the three RT measures for each *S* in an experimental condition, and this gave the basic RT scores on which analyses were performed.

Table 3 gives the group algebraic mean for each of the three RT scores in each experimental condition. There were no Error or Error Correction RT scores for certain events because *S* had only two choices when a certain event appeared: (*a*) either leave the control where it was, or (*b*) move it to the correct position. All *S*s always did the latter. All mean RT scores are negative, signifying that on the average they tended to occur after the onset of the stimulus, and this is the same tendency found by Adams and Xhignesse (1960) for two-dimensional, all-visual discrete tracking. The rank order of means in Table 3 for Correct RT scores is about the same as time on target and it is inferred from this that Correct RT measures, which made up most of the RT values at this advanced stage of practice, were principal determiners of the level of time-on-target curves. Mean differences between groups for Correct RT scores were evaluated by the *t* test for independent measures, and,

TABLE 3

GROUP ALGEBRAIC MEAN RT VALUES IN MILLISECONDS FOR THE THREE TYPES OF RESPONSES MADE IN EACH EXPERIMENTAL CONDITION: POOLED TRIALS 56 AND 57

Experimental Condition	Types of Responses		
	Correct RT	Error RT	Error Correction RT
CA	− 85	—	—
CAv	− 97	—	—
CV	−175	—	—
CaV	− 98	—	—
UA	−272	−264	−520
UAv	−364	−412	−780
UV	−356	−322	−614
UaV	−357	−403	−729

as with time on target, response to visual stimuli under bisensory conditions was superior to that in the unisensory visual task. The null hypothesis test of the mean difference in Correct RT scores for Cond. CV and CaV showed CaV to be significantly faster ($.02 < P < .05$). For uncertain events Cond. UaV had a performance level significantly below that of Cond. UV when time on target was the measure, but the comparison lacked significance at the .05 level when the measure was Correct RT score. Apparently this was because *S*s under Cond. UaV had a tendency to make more errors and take longer to correct each of them, and this produced a greater separation of time-on-target means. The mean frequency of errors for *S*s under Cond. UaV was 4.9, but only 1.50 for Cond. UV ($.05 < P < .10$). Table 3 shows that *S*s under Cond. UaV had a longer mean Error Correction RT than UV, but the mean difference was not significant either ($.30 < P < .40$). Despite the lack of statistical support, it seems reasonable to believe that the combined effects of these two error tendencies served to produce mean differences in the time-on-target measure that were absent when Correct RT scores were evaluated.

The mean difference between Correct RT scores for Cond. UA and UAv, like time on target, had superior performance for UA ($P < .01$). An added reason for the superiority of UA with time on target was that *S*s under Cond. UAv tended to make more errors and took longer to correct them. The mean frequency of errors for Cond. UAv was 10.1, and was 7.2 for UA, but this difference lacked significance at the .05 level. But a t test of the mean difference between Error Correction RT scores of Cond. UAv and UA showed that the tendency to take longer to correct an error under Cond. UAv was statistically significant ($P < .01$).

Mean Error RT scores in Table 3 have the same general ordering as Correct RT scores and t tests for related measures between mean Correct RT and mean Error RT scores for each of the four uncertain conditions produced no significant differences. When *S* made an initial response to a stimulus, whether the response was right or wrong, the RT was about the same and there was no basis for the a priori possibility that wrong responses might have different characteristics than correct responses.

Ordering of response pairs.—The greater speed of response to auditory than to visual stimulus has been known for a long time (Teichner, 1954), and this was true for the present study too. For Correct RT scores, the mean difference between Cond. CA and CV, and UA and UV, favored auditory events and a t test of each was significant ($P < .01$). Certain audio events in the bisensory task (CAv) induced about the same mean level of Correct RT score as the unisensory certain auditory events (CA) but, interestingly, the ordinarily slower response to visual events was speeded up in the bisensory task so that the mean of Correct RT scores for Cond. CaV is about the same as that for Cond. CA and CAv. Conversely, for uncertain events, Cond. UAv had performance significantly slower ($P < .01$) than the UA unisensory control in terms of Correct RT scores, and had about the same mean Correct RT score as Cond. UV and UaV.

The common mean speed of responding for a response pair in the bisensory task, with the speed determined by the predictability of the stimulus event pair, suggested a synchronization of the two responses. To check this, the recorder charts were evaluated for the order of each pair

TABLE 4

FOR GROUP AV, MEAN PERCENTAGE OF THE THREE POSSIBLE ORDERS OF THE TWO RESPONSES TO CERTAIN AND UNCERTAIN STIMULUS EVENTS

Response Order	Type of Stimulus Event Pair	
	Certain	Uncertain
Visual before auditory	10(0)	16(2)
Auditory before visual	11(0)	10(7)
Visual and auditory simultaneous	79(0)	46(19)

Note.—Primary entry is when both responses are correct. The value in parenthesis is the percentage when either one or both of the responses of the pair are initially in error. Based on 50 certain and 50 uncertain event pairs given each *S* on pooled Trials 56 and 57.

of responses to a simultaneous pair of stimuli. One response could precede the other, or they could occur together, and the criterion of togetherness used was a difference between RTs for the two responses that was in the range ±33 msec. When the difference was larger than this, the responding was considered sequential rather than simultaneous. Table 4 gives the mean percentages of the three possible orders of response for Group AV on Trials 56 and 57. Percentages are given for certain and uncertain stimulus event pairs when both responses are correct, and when one or both are initially in error. The predominant tendency was for simultaneous response movements, with the tendency being somewhat stronger for certain events. Given this joint responding in the bisensory task, it is concluded that the response pair was geared to the faster auditory stimulus when events were certain, and to the slower visual stimulus when events were uncertain.

Anticipatory responding.—Table 3 shows that the mean Correct RT scores of less than 100 msec. for Cond. CA, CAv, and CaV, are faster than the most optimally expected reaction time values (e.g., Klemmer, 1956; 1957) and therefore it was reasonable to infer that *S*s were predicting events rather than waiting for the stimulus to occur before responding. Helson (1949) was perhaps the first to observe that learned anticipation can enhance tracking proficiency by resulting in RTs less than optimum reaction time values. Bartlett (1951) and Adams (1961) discuss anticipation as a salient feature of skilled performance, and Poulton (1952, 1957) has introduced the term *perceptual anticipation* to describe the phenomenon. Recognizing that anticipation can result in excessively premature responding, with performance poorer than if *S* actually waited for the stimulus to occur, Adams and Xhignesse (1960) discuss *beneficial anticipation* where RTs result in less off-target time than expected on the basis of optimum reaction time. They further distinguish *perfect anticipation* where there is coincidence of stimulus and response occurrence.

In this study a Beneficial Anticipation RT was defined as being equal to or less than ±133 msec., and a Perfect Anticipation RT as being equal to or less than ±33 msec., which allowed for some minor deviation from exact coincidence in deference to equipment and measurement vari-

TABLE 5

MEAN PERCENTAGE OF RESPONSE TIME VALUES OBTAINED FOR CERTAIN EVENTS ON TRIALS 56 AND 57 THAT WERE BENEFICIALLY ANTICIPATORY OR PERFECTLY ANTICIPATORY

Experimental Condition	Beneficial Anticipation RT	Perfect Anticipation RT
CA	57%	24%
CV	33%	15%
CaV	56%	25%
CAv	57%	25%

ability. Table 5 gives the percentages of RTs for the certain events on Trials 56 and 57 that meet these two criteria. The major finding was that responses made to certain auditory events (CA and CAv), or to visual events that accompany certain auditory events (CaV), had a higher percentage of both types of anticipatory responses. To test statistically the influence of auditory events to promote a higher level of anticipation, a Beneficial Anticipation RT score was defined for an *S* as the number of Correct RTs that fell within the ±133 msec. criterion, and a Perfect Anticipation RT score as the number of Correct RTs that fell within the ±33 msec. range. It was found for the Beneficial Anticipation RT score that (*a*) Cond. CA was significantly higher than CV ($.01 < P < .05$), (*b*) the amount of anticipation for audio in the bisensory task (CAv) was not significantly different from the unisensory audio task (CA), and (*c*) Cond. CaV has a mean score that does not differ significantly from that of the bisensory audio condition, CAv. From all of this Cond. CaV would be expected to have a significantly higher Beneficial Anticipation RT score than Cond. CV, and this was the case ($.01 < P < .05$). These same tests for the Perfect Anticipation RT score gave the same results except that the mean difference between Cond. CA and CV lacked significance ($P > .05$). The inference from these comparisons is that anticipation is the mechanism that speeds the ordinarily slow response to the visual stimulus so that it occurs simultaneously with the faster response to the audio stimulus.

Discussion

The question we posed in the beginning does not have a single answer. Rather than finding an answer simply in terms of impairment or no impairment for response in a bisensory task, we found that response to certain visual events in the bisensory task was actually superior to the unisensory visual control group because RTs were increased to that of the faster response made to accompanying auditory stimuli. This gain for the visual modality was because the responses to certain visual-auditory stimulus pairs were usually made together and, through the potent behavioral mechanism of learned anticipation, the ordinarily slower visual RT increased to the same speed as response to auditory stimuli. Anticipation, being a learned and internal mechanism, triggered the response from within *S* and freed it from the constraints of the S-R loop via the visual system. Somewhat unexpectedly, we uncovered a set of circumstances where performance in a complex task was better than the aggregate performances of the part tasks comprising the whole.

On the other hand, when events were uncertain there was impairment because the ordinarily faster response to auditory stimuli was slowed down to occur simultaneously with the slower response to the visual stimulus. When stimulus uncertainty existed and the event had to be sensed before response could be made, anticipation was absent and the slower response governed the speed of the two response movements made together. One other item of impairment for audio response to uncertain bisensory stimuli was a significantly longer time to correct errors than for audio response to uncertain unisensory stimuli. This error phenomenon was not found for visual stimuli. Apparently this was because auditory errors were not as simply coded, but why correction responses for less than optimally coded errors should be subject to decoding delay in a bisensory task must await further research. Whatever the reason, it is another line of evidence showing that impairment in doing two things at once occurs when bisensory stimulus events are uncertain, and is in contrast to facili-

tating response effects when events are certain. The relevance of stimulus predictability for doing two things at once is consistent with the study of Bahrick and Shelly (1958).

One final aspect of our data deserves comment. Why did certain unisensory auditory events have a significantly higher level of beneficial and perfect anticipation than certain unisensory visual events? We can only surmise that audio events, being more difficult than visual in the early trials, produced less "rewarding" on-target time and so anticipation emerged as a learned mechanism that allowed *S* to achieve a satisfactory measure of success. This interpretation suggests the hypothesis that one of the variables determining anticipation is difficulty level of the task, in addition to regularities in the type of events and the times of their occurrence. A study by Adams and Xhignesse (1960) is consistent with this notion because they found anticipation in a two-dimensional, all-visual discrete tracking task to occur for more difficult 1-sec. events, but not for 2-sec. events. Using voluntaristic phrasing, it is as if *S* rises to the small crisis of not succeeding well enough and invokes anticipation as a behavioral device to solve his difficulties.

Summary

An experiment was performed to answer the question of whether a human *S* could do two things at once without impairment. A bisensory discrete tracking task was used where a probabilistic series of simultaneous auditory and visual stimuli were presented, each stimulus series for response with a separate hand. An auditory and a visual control group each practiced only a unisensory version of the task where response was with one hand. All events were of 2-sec. duration and had time certainty, but the type of event occurring next in the series could be either certain or uncertain.

The results revealed a net superiority of bisensory over unisensory responding when stimulus events were certain. This was because *S* in the bisensory task usually made the two response movements together, and anticipation of certain events resulted in an increase in speed of the visual response time to that of the faster audio response time. But, when events were uncertain, impairment was inferred for bisensory responding because the faster audio response time was reduced in speed and synchronized with the slower visual response time.

References

Adams, J. A. Human tracking behavior. *Psychol. Bull.*, 1961, **58**, 55–79.

Adams, J. A., & Xhignesse, L. V. Some determinants of two-dimensional visual tracking behavior. *J. exp. Psychol.*, 1960, **60**, 391–403.

Attneave, F. *Applications of information theory to psychology.* New York: Holt, 1959.

Bahrick, H. P., & Shelly, C. Time sharing as an index of automatization. *J. exp. Psychol.*, 1958, **56**, 288–293.

Bartlett, F. C. Anticipation in human performance. In G. Ekman, T. Husén, G. Johansson, & C. I. Sandström (Eds.), *Essays in psychology.* Uppsala: Almquist & Wiksells, 1951. Pp. 1–17.

Bricker, P. D. Information measurement and reaction time: A review. In H. Quastler (Ed.), *Information theory in psychology.* Glencoe: Free Press, 1955. Pp. 350–359.

Helson, H. Design of equipment and optimal human operation. *Amer. J. Psychol.*, 1949, **62**, 473–497.

Klemmer, E. T. Time uncertainty in simple reaction time. *J. exp. Psychol.*, 1956, **51**, 179–184.

Klemmer, E. T. Simple reaction time as a function of time uncertainty. *J. exp. Psychol.*, 1957, **54**, 195–200.

Lindquist, E. F. *Design and analysis of experiments in psychology and education.* New York: Houghton Mifflin, 1953.

Poulton, E. C. Perceptual anticipation in tracking with two-pointer and one-pointer displays. *Brit. J. Psychol.*, 1952, **43**, 222–229.

Poulton, E. C. On prediction in skilled movements. *Psychol. Bull.*, 1957, **54**, 467–478.

Teichner, W. H. Recent studies in simple reaction time. *Psychol. Bull.*, 1954, **51**, 128–149.

(Received December 2, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 2, 207–213

NONREINFORCEMENT AND NEUTRALIZATION OF STIMULI [1]

DAVID LABERGE, JAMES G. GREENO,[2] AND O. F. PETERSON

University of Minnesota

A problem common to many learning situations is the determination of the effects of nonreinforcement. It is commonly inferred that a reinforced trial strengthens the response made, although there may remain some question as to whether or not this involves also a weakening of other responses. A nonreinforced trial, on the other hand, gives rise to many conflicting interpretations, including (*a*) weakening of the response made, (*b*) strengthening of a competing response, (*c*) no effect on responses, and even (*d*) strengthening of the response made.

When applied in the two-choice prediction situation, where *S* predicts event E_1 by making response A_1 and predicts E_2 by making response A_2, the nonreinforcing event, labeled E_0, occurs when neither E_1 nor E_2 follows the response. Results of some studies (Atkinson, 1956; Neimark, 1956) designed to test theoretical interpretations of nonreinforcement in this situation have supported the identity hypothesis, namely that the average effect of a nonreinforced trial is to leave the response probabilities unchanged. Anderson and Grant (1957) have reported results which are reasonably close to predictions on the basis of the identity hypothesis, but the fact that their sharing parameter for E_0 trials apparently did not reach unity prevented a clear confirmation of the identity hypothesis. They also showed a tendency for *S*s to repeat on the next trial the response which was followed by an E_0 event.

It has been shown (Estes, 1959) that the process underlying the identity operation may be (*a*) a strengthening of the response just made through conditioning by contiguity, or (*b*) a weighted combination of the weakening of the response plus strengthening of alternative responses, or (*c*) no effect. The chief concern of this study is to investigate an additional process which could result from nonreinforcement.

Experimenters commonly observe that the introduction of nonreinforcement occasions *S* to make responses in addition to those recorded by *E*. If these responses were reinforced in some way, then presumably some stimulus elements would become connected to them, and these elements would then be regarded as neutral with respect to the A_1 and A_2 responses. In terms of statistical learning theory, therefore, a nonreinforced trial might well result in the neutralization of some of the sampled choice-point cues. If in a random sample all of the cues were neutralized, then the average relative proportion of choice cues would remain constant, and no change in the response probabilities would occur on the next trial. If only a part of the sample were neutralized, the net effect on subsequent response probability would depend upon what happened to the rest of the sampled cues.

This study attempted to investigate

[1] This study was supported by grants to the senior author from the National Science Foundation and from the Graduate School of the University of Minnesota.

[2] Now at Indiana University.

the neutralization of cues in the simple noncontingent prediction situation. The task of *S* was to predict which of two lights would follow a signal. One possible instance of the neutral response class here might be *S*'s implicit prediction that neither light would occur following the signal, even though he was required to make one of the two overt responses associated with the lights. And presumably when neither light flashed, this might have reinforced the implicit "neither light" prediction.

There were two parts in the present investigation. The first part involved comparisons of slopes of learning curves under a reinforcement ratio of 90:10, following differential training at a reinforcement ratio of 50:50. It was expected from the deductions described in an earlier paper (LaBerge, 1959b), that groups having a higher percentage of nonreinforced trials during 50:50 training should have a greater proportion of neutral elements in the population at the start of 90:10 training, and therefore should show faster rates of learning during the 90:10 schedule.

The second part of this study involved a comparison of learning performance among groups having different numbers of nonreinforced trials inserted in a 90:10 schedule. In this case, the neutralization hypothesis implies conflicting predictions, depending upon what other assumptions are made. If the average sample contains representative proportions of choice-point cues, and if the whole sample is neutralized by a nonreinforced trial, then groups with the higher proportions of nonreinforced trials should show higher learning rates. If a part of the sample is neutralized and a part connected to the alternative response, then groups with the higher proportions of nonreinforced trials should show lower learning rates.

METHOD

Apparatus.—Each of the two booths in a darkened room contained a response box with two spring-loaded levers and a reinforcing light above each lever. The buzzer signal and reinforcing lights were controlled by a timer, and responses were recorded automatically.

Design.—There were two main parts in the experiment, three groups in each part, and 48 *S*s per group. Both parts were run concurrently. In Part I, groups were given 0, 33, or 67% E_0 trials by random combination of 0, 30, or 120 E_0 trials with 60 reinforced trials having an E_1 to E_2 ratio of 50:50. Following this, *S*s were immediately shifted to a reinforcement schedule of 90:10 without E_0 events for 100 trials. In Part II, groups were given 0, 33, or 67% E_0 trials by combination of 0, 50, or 200 E_0 trials with 100 reinforced trials having an E_1 to E_2 ratio of 90:10. Sequences of reinforcing and nonreinforcing events were constructed by obtaining 20-trial blocks of reinforced trials and randomly inserting in these appropriate numbers of E_0 trials. Each group had 12 different noncontingent sequences, 4 *S*s to a sequence, and sequences were matched wherever possible across all groups. Side of the more frequent reinforcing light was counterbalanced within each sequence subgroup.

Subjects.—The 288 *S*s were volunteers from introductory psychology courses at the University of Minnesota.

Procedure.—The standard prediction instructions (Estes & Straughan, 1954) were given with no interpretation for the nonreinforced trials. Practice trials were presented with outcomes in the order $E_1E_1E_2E_2$-$E_0E_0E_1E_2$. If there were questions concerning the E_0 trials, *E* replied, "That may happen on some trials. In any case, make a choice on each trial."

The sequence of trials was then given without interruptions. The timing of events within trials was as follows: Duration of the signal, 2 sec.; time between the signal and the reinforcing light, 1 sec.; duration of the reinforcing light, 1 sec.; time between the reinforcing light and the next signal, 2 sec.

RESULTS

Part I.—Shown in Fig. 1 are the curves of learning under a reinforce-

ment ratio of 90:10, following three different conditions of preliminary training. Analysis of learning rates was based on the mean number of A_1 responses over the 100 trials, yielding values of 84.58, 83.25, and 80.97 for the groups having 67%, 33%, and 0% nonreinforcements, respectively, during preliminary trials. The linear trend among these means yielded a significant F (4.97; $df = 1/72$; $P < .05$) in the analysis of variance for response curves (Ostle, 1954).

In order to investigate the effect of E_0 trials on the rate of learning in a more detailed fashion, a statistic, ϕ, was used, defined by:

$$\phi = \frac{\sum_{i=1}^{n'} (p_{i+1} - p_i)}{\sum_{i=1}^{n'} (1 - p_i)}$$

counting only the trials on which E_1 was given on Trial i. This statistic estimates θ in the model presented by Estes and Straughan (1954). In the present model, ϕ is greater than θ whenever the proportion of neutral elements is greater than zero. It is therefore predicted that the groups receiving E_0 trials in the pretraining sequence should show higher values of ϕ than the group receiving no E_0 trials. Block-by-block estimates of ϕ for these groups are graphed in Fig. 2. Estimation of ϕ for each S over all trials yielded means of .087, .135, and .163 for the 0%, 33%, and 67% groups, respectively. A variance estimate of .0321 (df = 72) was obtained as the residual MS of the factorial analysis of variance, and 95% confidence intervals were estimated as follows:

$$\tfrac{1}{2}(\bar{\phi}_{33} + \bar{\phi}_{67}) - \bar{\phi}_0 = .062 \pm .063;$$
$$\bar{\phi}_{67} - \bar{\phi}_{33} = .028 \pm .073$$

The method of orthogonal contrasts thus revealed a difference of debatable significance between the no-E_0 group and the E_0 groups combined, whereas the difference between the

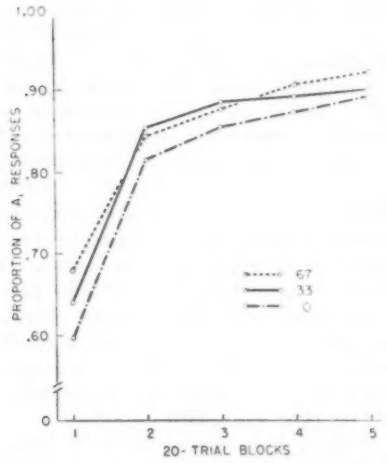

FIG. 1. Learning curves under a 90:10 reinforcement ratio with no E_0 trials, following training under a 50:50 reinforcement ratio with 0%, 33%, or 67% E_0 trials.

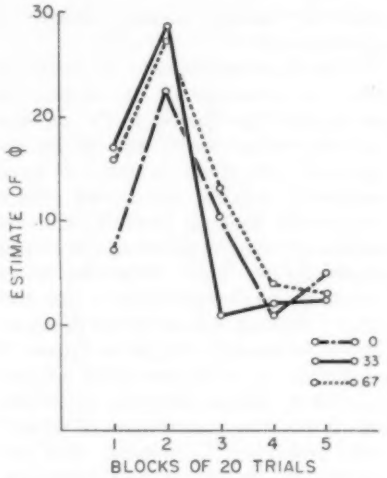

FIG. 2. Block-by-block estimates of ϕ for the 90:10 reinforcement condition of Part I.

two E_0 groups clearly failed of significance.

The data presented in Table 1 allow an estimate of *Ss*' tendencies to repeat the choice made on the previous trial during the 50:50 sequences. Since the two choices were reinforced equally often, the choice frequencies include both A_1 and A_2 responses, and A_i refers to the choice reinforced on Trial n for the trials following reinforced trials. For the trials following E_0 trials, the designation is arbitrary. Since no choice is reinforced by an E_0 event, the proportions of A_i choices following A_i and E_0 are simply the proportions of repetitions following E_0 events, and the proportions of A_i choices following A_j and E_0 are the proportions of alternations following E_0 events. In this study, as in the study of Anderson and Grant (1957), *Ss* showed a higher proportion of A_i choices following E_0 trials on which they chose A_i than they did following E_0 trials on which they chose A_j. As Table 1 also indicates, *Ss* tended to repeat their previous choices on trials following reinforced trials as well.

Part II.—The effect of the proportion of nonreinforced trials during learning under a 90:10 reinforcement ratio is shown in Fig. 3. Each point on the curves represents a block of trials which includes 20 reinforcements. The mean percentages of A_1 responses over all trials were 84.38, 78.79, and 75.67 for groups having 0%, 33%, and 67% nonreinforcements, respectively. The linear trend among these means was significant ($F = 24.34$; $df = 1/72$, $P < .0005$) in the analysis of variance for response curves (Ostle, 1954). Higher proportions of A_1 responses were associated with the groups having the least percentage of nonreinforced trials. There is also an apparent separation of group asymptotes, although a longer series of trials is needed to establish this conclusion firmly.

Figure 4 shows the block-by-block estimates of ϕ for these groups. The means for groups receiving 0%, 33%, and 67% E_0 trials were .108, .324, and .247, respectively. The residual *MS*

TABLE 1

PROPORTION OF A_i RESPONSES ON TRIAL $n+1$ UNDER THE 50:50 CONDITION OF PART I

% E_0	Response on Trial n	Block 1		Block 2		Block 3	
		E_{in}	E_{on}	E_{in}	E_{on}	E_{in}	E_{on}
0	A_{in}	.50	—	.61	—	.60	—
	A_{jn}	.42	—	.46	—	.53	—
33	A_{in}	.50	.50	.60	.53	.61	.54
	A_{jn}	.42	.50	.55	.47	.52	.46
67	A_{in}	.52	.53	.61	.59	.59	.54
	A_{jn}	.48	.47	.50	.41	.55	.46

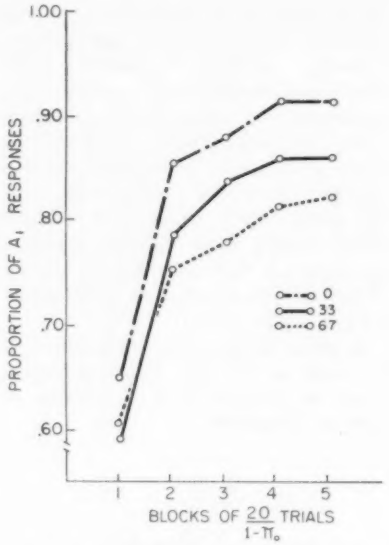

FIG. 3. Learning curves under a 90:10 reinforcement ratio with 0%, 33%, or 67% E_0 trials.

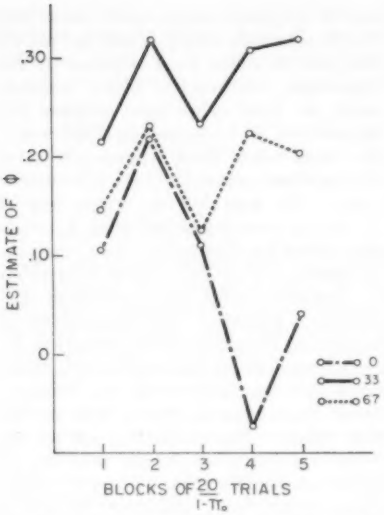

FIG. 4. Block-by-block estimates of ϕ for the 90:10 reinforcement condition of Part II. (The groups were given 0%, 33%, or 67% E_0 trials.)

estimate of σ^2 from the factorial analysis of variance was .0546 ($df = 72$) and 95% confidence intervals were estimated as follows:

$$\tfrac{1}{2}(\bar{\phi}_{33} + \bar{\phi}_{67}) - \phi_0 = .177 \pm .082;$$
$$\bar{\phi}_{67} - \bar{\phi}_{33} = -.077 \pm .096$$

Thus, orthogonal contrasts revealed a significant difference between the no-E_0 group and the E_0 groups combined, whereas the difference between the two E_0 groups again failed of significance.

Table 2 contains the results of an analysis similar to that presented in Table 1. However, the lack of symmetry under the 90:10 reinforcement schedules requires a somewhat different presentation. The columns labeled "E_{in}" in Table 2 contain proportions of choices on trials following reinforced trials, with E_1 and E_2 trials pooled. The partition involving A_{1n} and A_{2n} does not, of course, involve pooling over the responses on Trial n as is the case in Table 1. Since the proportions of A_1 choices are consistently higher following A_{1n} than following A_{2n}, it appears that *Ss* tended to repeat their previous choices following both E_0 and E_i trials, as they did under 50:50 training of Part I.

DISCUSSION

The main hypothesis under test was that nonreinforcement connects cues to responses other than the choice alternatives in a two-choice learning situation. This hypothesis was supported in the first part of the experiment where groups having a greater proportion of E_0 trials showed higher rates of learning on a

TABLE 2

PROPORTION OF A_1 RESPONSES ON TRIAL $n + 1$ UNDER THE 90:10 CONDITION OF PART II

% E_0	Response on Trial n	Block 1		Block 2		Block 3		Block 4		Block 5	
		E_{in}	E_{on}	E_{in}	E_{on}	E_{in}	E_{on}	E_{in}	E_{on}	E_{in}	E_{on}
0	A_{1n}	.79	—	.89	—	.91	—	.94	—	.92	—
	A_{2n}	.42	—	.64	—	.70	—	.69	—	.79	—
33	A_{1n}	.78	.51	.88	.75	.92	.80	.92	.81	.91	.83
	A_{2n}	.53	.41	.66	.55	.63	.56	.78	.58	.80	.74
67	A_{1n}	.72	.68	.84	.81	.86	.85	.90	.87	.87	.87
	A_{2n}	.55	.46	.64	.52	.64	.56	.65	.62	.71	.56

subsequent 90:10 reinforcement condition. This result and the theoretical background are strengthened considerably by comparing the present group orderings with those in a related study (LaBerge, 1959b), in which neutral elements were depleted by 50:50 training with no E_0 events. In the present study, groups having longer preliminary training showed faster learning, whereas in the previous study groups having longer preliminary training showed slower learning. Both results are consistent with the neutral elements assumption.

Quantitative analysis in terms of the neutral elements model yields some additional support for the neutralizing hypothesis. If the probability of an E_0 event during the 50:50 trials is .67, and if it is assumed that all the sample cues are neutralized on an E_0 trial, then it is expected by probability matching considerations that the population of cues at the beginning of the 90:10 condition will contain 67% neutral elements. When we compare the 67% E_0 group of Part I with the 0% group of Part II, we note that their curves are virtually identical. An independent estimate of the proportion of neutral elements at the beginning of training for the 0% group of Part II constitutes a check on the neutral elements proportion of .67. The relevant parameters estimated for a similar study (LaBerge, 1959a) were .755 for the initial proportion of neutral elements, and .032 for θ. With these parameter values, and allowing for the practice trials in the present experiment, we obtain .66 for the estimated proportion of neutral elements at the beginning of the 90:10 condition for the 0% group of Part II. Thus it would seem that a schedule of 67% E_0 events maintained the proportion of neutral elements at its original level during the 50:50 schedule.

However, it would seem theoretically reckless to let the close quantitative agreement between these values rule out possible additional processes during E_0 trials. A likely candidate would be a correction factor, namely the strengthening of the response opposite to the one made. This could account for the ordering of response frequencies under the 90:10 sequences with E_0 trials in Part II. The data in Table 2 are relevant to this hypothesis, since for a given response made on Trial n we can compare the proportions of A_1 choices on Trial $n + 1$ for times when Trial n was reinforced and for times when Trial n was not reinforced. We note that the proportions of A_1 choices were lower following E_0 trials than following reinforced trials when A_1 was performed on Trial n, a finding which is consistent with the correction hypothesis. However, the proportions of A_1 choices were also lower following E_0 trials than following reinforced trials when A_2 was performed on Trial n. These data suggest, then, that an E_0 trial reduces the probability of an A_1 choice regardless of which response is chosen on the E_0 trial. The correction effect shown here, therefore, is not a "pure" one, for it apparently occurs only when the more frequent response is made on an E_0 trial.

Another way of evaluating the evidence for a general correction effect consists of examining the differences among the groups shown in Fig. 3. If E_0 did not affect response probabilities, as expected from an identity operation, then the three group curves should be identical, since the groups differed only with respect to the presence of E_0 trials which were added to equivalent series of reinforced trials. The fact that instead we obtained a lowering of A_1 proportions with increasing proportions of E_0 trials argues for some sort of average correction effect, while seriously questioning the validity of the identity hypothesis.

Since slopes of response probability curves are rather insensitive measures of the proportion of neutral elements in the population, it is expected that even a small correction factor in the E_0 process would easily mask any effect on learning rate due to neutralizing. Thus, it should be no surprise to find that increased proportions of E_0 trials during the 90:10 condition lowered the level of A_1 responding as shown in Fig. 3. If a neutralization effect did occur over these

trials, it should be revealed by the ϕ curves of Fig. 4, which measure rate of learning due to E_1 reinforcements. The slopes of these curves are apparently less steep for groups with E_0 trials, which suggests that rate of learning for these groups is higher for a longer time, compared with the group having no E_0 trials.

Summary

Six groups of 48 *Ss* each were trained in a noncontingent two-choice situation in an attempt to investigate effects of nonreinforcement upon learning. The main hypothesis tested was that nonreinforcement connects at least part of the sampled elements to responses other than the choice alternatives. Three of the groups were given preliminary trials under a 50:50 reinforcement ratio with 0%, 33%, or 67% of the trials nonreinforced. These groups were then shifted to a 90:10 reinforcement condition without further nonreinforcement. Groups having greater amounts of nonreinforcement during preliminary training showed significantly higher rates of learning under the 90:10 condition, supporting the neutralization hypothesis.

For further analysis of the neutralization effect and in order to test the identity hypothesis, 0%, 33%, or 67% of nonreinforced trials were inserted in a 90:10 reinforcement schedule. Groups having more nonreinforced trials showed lower proportions of A_1 responses, indicating that nonreinforcement is not in general a neutral event, as required by the identity hypothesis. Estimates of learning rates by a conditional probability method supported the neutralization hypothesis for these groups.

References

Anderson, N. H., & Grant, D. A. A test of a statistical learning theory model for two-choice behavior with double stimulus events. *J. exp. Psychol.*, 1957, **54**, 305–317.

Atkinson, R. C. An analysis of the effect of nonreinforced trials in terms of statistical learning theory. *J. exp. Psychol.*, 1956, **52**, 28–32.

Estes, W. K. The statistical approach to learning theory. In S. Koch, (Ed.), *Psychology: A study of a science.* Vol. 2. New York: McGraw-Hill, 1959.

Estes, W. K., & Straughan, J. H. Analysis of a verbal conditioning situation in terms of statistical learning theory. *J. exp. Psychol.*, 1954, **47**, 225–234.

LaBerge, D. Effect of preliminary trials on rate of conditioning in a simple prediction situation. *J. exp. Psychol.*, 1959, **57**, 20–24. (a)

LaBerge, D. A model with neutral elements. In R. R. Bush & W. K. Estes (Eds.), *Studies in mathematical learning theory.* Stanford: Stanford Univer. Press, 1959. (b)

Neimark, E. D. Effects of type of nonreinforcement and number of alternative responses in two verbal conditioning situations. *J. exp. Psychol.*, 1956, **52**, 249–252.

Ostle, B. *Statistics in research.* Ames, Iowa: Iowa State Coll., Press, 1954.

(Received December 6, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 2, 214–215

SUPPLEMENTARY REPORT: PARTIAL REINFORCEMENT AND AMOUNT OF REINFORCEMENT AS DETERMINANTS OF INSTRUMENTAL LICKING RATES[1]

STEWART H. HULSE AND W. EDWARD BACON

Johns Hopkins University

Hulse (1958) used large and small wet-mash rewards and examined resistance to extinction of a running response after these amounts had been given on 50% and 100% of the training trials. After partial reinforcement, the large reward produced much faster running than the small reward, but after continuous reinforcement, the large reward produced slower running than the small reward. The present experiment generalizes this observation to instrumentally-conditioned licking, following a suggestion (Hulse, Snyder, & Bacon, 1960) that the phenomenon might hold for licking as well as for running.

Method.—The *S*s were 50 male albino rats. None had prior experience in licking from tubes. The apparatus and the general technique for instrumentally conditioning licking have been described elsewhere (Hulse, 1960; Hulse et al., 1960). In the present experiment, two schedules of reinforcement were used: 8 to 1 fixed ratio (FR8) and continuous reinforcement (CRF). These schedules were combined in a factorial design with five concentrations of a saccharine reinforcer: .25, .50, 1.80, 4.50, and 10.00 gm. of saccharine added to 1 liter of tap water. Each reinforcement consisted of a .005-cc drop of the appropriate concentration of saccharine. The *S*s were tested both hungry and thirsty. Food (7.5 gm.) and water (20 cc minus the amount obtained in the experimental apparatus) were given in the form of wet mash .5 hr. to 1.0 hr. after testing on each day.

On Day 1 of training, all *S*s were permitted 100 licks of the appropriate concentration of saccharine on a CRF schedule. Licking was then shaped such that by Day 7, all *S*s were producing 800 licks on the appropriate schedule of reinforcement. This regime was continued through Day 20. On Day 21, 4 days of extinction were begun. Each day, *S*s were permitted to lick on the dry tube for 5 min.

Results.—During training, daily licking rates were computed from the times *S*s required to emit the first 400 and the last 400 licks. These data, averaged across the last 5 days of training, are shown in Fig. 1. Three general effects ($Ps < .001$ in an analysis of variance) are shown in the figure. First, licking rate increases monotonically as a function of the log of reinforcement concentration. Second, with CRF, rates decrease

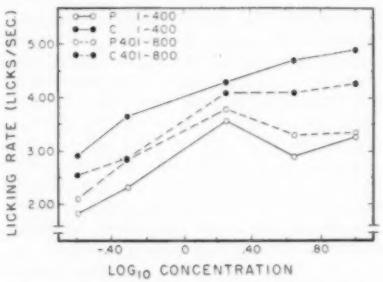

FIG. 1. Acquisition licking rates for FR8 and CRF (P and C, respectively) and five concentrations of saccharine. (Means are shown for the first and last 400 licks.)

from the first to the last half of a session, while with partial reinforcement, rates increase across the two blocks of licks. Finally, partial reinforcement yields slower average rates of licking than CRF. This last observation bears further investigation since, with responses other than licking, ratio reinforce-

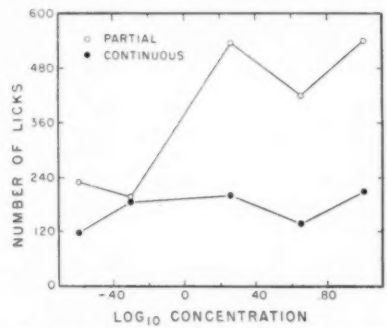

FIG. 2. Licks during extinction following FR8 and CRF with five concentrations of saccharine.

[1] This research was supported by National Science Foundation Research Grant G8712 to the senior author.

ment in a free operant situation yields a higher rate of responding than CRF.

During extinction, concentration, schedule, and their interaction are all significant ($Ps < .01$). Figure 2 shows that partial reinforcement produces greater resistance to extinction than CRF. Also, the number of licks emitted during extinction increases as a function of the log of reinforcement concentration. This latter relation is not simple, however, since by t test, none of the CRF groups differ significantly from one another ($Ps > .05$). With partial reinforcement, on the other hand, resistance to extinction increases markedly as a function of reinforcement concentration. By t test, the means for the three highest concentrations all differ significantly from those for the two lowest concentrations ($Ps < .01$).

References

HULSE, S. H., JR. Amount and percentage of reinforcement and duration of goal confinement in conditioning and extinction. *J. exp. Psychol.*, 1958, **56**, 48–57.

HULSE, S. H., JR. A precision liquid feeding system controlled by licking behavior. *J. exp. Anal. Behav.*, 1960, **3**, 1–3.

HULSE, S. H., SNYDER, H. L., & BACON, W. E. Instrumental licking behavior as a function of schedule, volume, and concentration of a saccharine reinforcer. *J. exp. Psychol.*, 1960, **60**, 359–364.

(Received February 4, 1961)

Journal of Experimental Psychology
1962, Vol. 63, No. 2, 215–216

SUPPLEMENTARY REPORT: SHOCK INTENSITY AND UNCONDITIONED RESPONDING IN A SHUTTLE BOX

THOMAS R. TRABASSO[1] AND RICHARD W. THOMPSON

Michigan State University

Kimble (1955) studied UCRs of rats to different intensities of shock. He found that shock elicited two kinds of responses, "flinching" at lower intensities (0.1–0.6 ma.) and "jumping" which increased in frequency as shock intensity increased (0.1–0.9 ma.). However, he used an ascending-descending order of shock presentation and the dimensions of his shockbox (11 in. long, 9 in. high, and 4¼ in. wide) differed from those of typical shuttleboxes (24–36 in. long and 5–14 in. high). Since the order of shock presentation may influence the frequency of responding and the dimensions of the shuttlebox have been found to limit the response repertoire of S (Denny & Thomas, 1960) the present study was designed to replicate Kimble's with modification of these variables.

Method.—The *Ss* were 18 male and 18 female albino rats (Kimble's *Ss* were 10 males), ranging in age from 94 to 169 days (Kimble's *Ss* were aged 200 days); groups matched for age and sex were randomly assigned to the two experimental conditions (ceiling heights).

The apparatus was a shuttlebox, 36 in. long and 4 in. wide with either a 5 in. or a 14 in. ceiling. The stimulator, timer, adaptation period, shock duration (1 sec.), intershock interval (30 sec.), and intensity range (0.0–0.9 ma. in 0.1 ma. steps) were identical to those used by Kimble. Two sessions of 80 shocks with a 5-min. rest interval between sessions were given. Each of the 10 levels of shock was given eight times in a different random order for each session and each S.

The responses were operationally defined: (*a*) No Response; (*b*) Flinch—S made a startle response to shock, no more than one paw leaving the grid; (*c*) Shuttle-Compatible—this response class is analogous to Kimble's "Jump" category and was divided, in the present study, into three subclasses Run, Prance—S advanced or remained within a 9-in. section with both hind feet leaving the grid nearly simultaneously and alternately raised one of the forefeet, and Jump—all four of S's feet left the grid. The main response classes (*a*, *b*, and *c*) are mutually exclusive but the subclasses of Class *c* are not.

Results and discussion.—For the Flinch and Shuttle-Compatible response classes a separate analysis of variance for each class, using shock levels as repeated measures was performed. For both classes, the only significant effect was attributed to shock intensity (Flinch, $F = 53.53$; Shuttle-Compatible, $F = 293.29$; $8/256$ df, $P < .01$). Neither ceiling height, sex, nor any interactions were significant in either analysis. The data for the 36 *Ss* were summarized for

[1] National Science Foundation Cooperative Fellow. The authors wish to thank M. Ray Denny for his helpful advice and criticism throughout the course of this study.

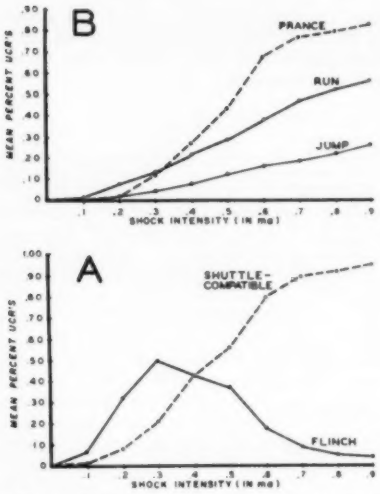

FIG. 1. Mean percentage of *Ss* showing (A) Flinch and Shuttle-Compatible, and (B) Prance, Run, and Jump responses as a function of shock intensity.

the Flinch and Shuttle-Compatible classes and are presented as mean percentages of responding as a function of shock intensity in Fig. 1A.

In general, our results conform very well to those reported by Kimble. The *Ss* in the present study flinched more frequently at all levels than Kimble's (e.g., Kimble's *Ss* flinched 36% at 0.3 ma. vs. 50% in Fig. 1A, at the same point). With respect to the Shuttle-Compatible responses, there were no apparent differences between Kimble's "Jump" data and ours.

Separate analyses of the Shuttle-Compatible subclasses (Run, Prance, and Jump) using the two sessions as repeated measures yielded no significant interactions or differences between sexes or ceiling heights. These data are summarized in Fig. 1B. The percentages reported do not add to 100% since it was possible for an individual *S* to give all three responses to one shock presentation.

Kimble reported a significant difference in median threshold for jumping between first and second halves of his experiment (.36 ma. vs. .51 ma.). There were no significant differences between our medians (.439 ma. vs. .468 ma.; $t = 1.18$, $df = 35$, $P > .05$). Neither did the frequency of the Flinch responses differ between the two sessions ($t = 0.61$). Hence, no evidence of adaptation to shock was observed.

The findings of the present study support the generality of Kimble's results. Presumably, identification of the UCRs to shock is necessary for a thorough understanding of escape and avoidance learning.

REFERENCES

DENNY, M. R., & THOMAS, J. O. Avoidance learning and relearning as a function of shuttlebox dimensions. *Science*, 1960, **132**, 620–621.

KIMBLE, G. A. Shock intensity and avoidance learning. *J. comp. physiol. Psychol.*, 1955, **48**, 281–284.

(Received March 18, 1961)